Itkowsky, F.A.

SENSITIVITY OF VARIABLE PROBABILITY SAMPLING ESTIMATES TO INITIAL LANDSAT CLASSIFICATION

Frank A. Itkowsky and James A. Smith

Final Report, Volume I

Rocky Mountain Forest and Range Experiment Station

U.S. Forest Service

Cooperative Agreement 16-741-CA

September, 1978

College of Forestry and Natural Resources

Colorado State University

Fort Collins, Colorado 80523

LIBRARY

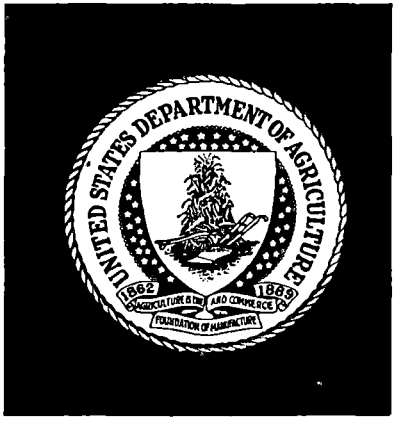
FEB 16 1979

ROCKY MOUNTAIN STATI

AD-33 Bookplate
(1—63)U. S. FOREST SERVICE

NATIONAL FT. COLLINS

AGRICULTURAL



LIBRARY

LIBRARY COPY
ROCKY MTN. FOREST & RANGE
EXPERIMENT STATION

SENSITIVITY OF VARIABLE PROBABILITY SAMPLING ESTIMATES TO INITIAL LANDSAT CLASSIFICATION

Frank A. Itkowsky and James A. Smith

Final Report, Volume I

Rocky Mountain Forest and Range Experiment Station

U.S. Forest Service

Cooperative Agreement 16-741-CA

September, 1978

College of Forestry and Natural Resources

Colorado State University

Fort Collins, Colorado 80523

LIBRARY COPY
ROCKY MTN. FOREST & RANGE
EXPERIMENT STATION

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL U.S. FOREST SERVICE POSITION, UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.

THE FOLLOWING REPORT IS FROM A THESIS SUBMITTED TO THE GRADUATE FACULTY OF COLORADO STATE UNIVERSITY IN PARTIAL FULFILLMENT OF DEGREE OF MASTER OF SCIENCE.

FOREWARD

The research described in this report was supported by the U.S. Forest Service under Cooperative Agreement CA-16-741, "Signature Studies for Natural Resource Inventory." It summarizes results obtained under Objective 1 of the project entitled, "Middle Park Inventory Analysis" and covers the period July 1, 1977 to August 30, 1978. Mr. Robert Dana was technical monitor of this project. The purpose of this study was to determine the effects of initial Landsat classifications on final acreage estimates obtained by employing variable probability sampling. Digital Landsat data and aerial photography for the Middle Park Basin of north-central Colorado were utilized.

Participating project personnel at various times included: Dr. James A. Smith, Principal Investigator; Mr. Frank Itkowsky, Graduate Research Assistant; Mr. K. Jon Ranson, Research Associate; Mr. Steve Bernath, Research Associate; Mr. Dan Kimes, Graduate Research Assistant; Ms. Julie Kirchner, Graduate Research Assistant; and Mr. Jim Kramer, Graduate Research Assistant.

Mr. Frank Itkowsky received the Master of Science degree for work related to project tasks and the bulk of this report is extracted from his thesis paper.

The authors would like to express their appreciation to Mr. Robert Dana, Mr. Dick Francis, Dr. Gary Dixon, and Mr. Ed Roberts, all of the Rocky Mountain Forest and Range Experiment Station; Dr. Brian Kent, Department of Forest and Wood Sciences, Colorado State University; and

Dr. David Bowden, Department of Statistics, Colorado State University for their helpful assistance. The authors particularly express their appreciation to Dr. Richard Driscoll, Program Leader, Rocky Mountain Forest and Range Experiment Station whose support made this work possible.

ABSTRACT

The Landsat satellite system provides digital data which can be processed to produce land cover classification maps over vast areas. A much greater degree of detail can be obtained using aerial photography, but the time consuming human interpretation required precludes its use over large areas. Sampling is used, therefore, in conjunction with aerial photography to provide detailed estimates of specific natural resources parameters. A variable probability sampling technique can be used which utilizes a Landsat land cover classification to obtain variable selection probabilities of sample units. This technique requires a fewer number of sample units to be interpreted on the aerial photography to produce a predetermined degree of accuracy, than simple random sampling.

The purpose of this study was to determine the effects of relative Landsat classification accuracy on variable probability sampling acreage estimates. If sample estimates are very sensitive to classification accuracy, it would behoove one to produce accurate classifications to obtain accurate estimates. If the converse is true, savings in human and computer time can be made in the classification procedure by producing classifications where a high degree of accuracy is not necessary.

The first step of this study was to produce a land cover classification of the Middle Park study area of north-central Colorado using

Landsat digital data. A supervised classification technique was employed. Four additional classifications of the same area were obtained using different techniques. Samples were then conducted to estimate the acreage of coniferous forest and meadow in the study area using each classification to obtain selection probabilities. Sampling was conducted on color infrared aerial photography at a scale of 1:24,000 to obtain detailed acreage measurements. Selection probabilities from Landsat and corresponding acreage estimates from aerial photography were then used in the sampling formulas to produce acreage estimates for the entire study area. Estimates obtained using the different classifications were compared to determine the effects, if any, caused by the classification being used.

It was found that meadow acreage estimtes of sufficient confidence could not be obtained without sampling an inordinately large portion of the population. This was attributed to the nature of this land cover class rather than the sampling procedure used. Several samples were conducted to produce conifer acreage estimates. It was apparent that greater differences in estimates and variances of the estimates were caused by the specific sample units selected than by the Landsat classification used. A final sample selecting 19 out of a possible 100 sample units was then conducted and a t test run which showed that the estimates were not significantly different at the 95% confidence level. It was thus concluded that, in this case, Landsat classification accuracy did not have a significant effect on sample estimates.

TABLE OF CONTENTS

Chapter		Page
1.0	INTRODUCTION	1
	1.1 Study Objective	2
	1.2 Experimental Procedure	4
2.0	PROCEDURE FOR PRODUCING THE LAND COVER CLASSIFICATIONS OF MIDDLE PARK	6
	2.1 Study Site Description	7
	2.2 The ORIGINAL Classification	9
	2.2.1 Data Preprocessing	10
	2.2.2 Class Determination/Description	10
	2.2.3 Signature Extration	14
	2.2.4 Classification and Display	22
	2.3 The JOSESIGS Classification	23
	2.4 The SCRAMBL1, SCRAMBL2, and SCRAMBL3 Classifications	23
	2.5 Relative Classification Performance	27
3.0	MULTISTAGE AND VARIABLE PROBABILITY SAMPLING	32
4.0	DESCRIPTION OF VARIABLE PROBABILITY SAMPLING EXPERIMENT	47
	4.1 Selection of Sample Unit Size and Shape	48
	4.2 List Sampling Procedure	50
	4.3 Obtaining Sample Measurements from Aerial Photography	59
	4.4 Calculations and Results	63
5.0	CONCLUSIONS	93
LITERATURE CITED		98
APPENDIX A	Confusion Matrices showing Classification Accuracy	100
APPENDIX B	Sample Lists	106
APPENDIX C	Sample Data	127
APPENDIX D	PSU AND SCRAMBL Computer Program Lists	159
APPENDIX E	The Use of Variable Probability Sampling in Landsat Classification Accuracy Estimation	179

LIST OF TABLES

Table		Page
1	Confusion matrix used to produce the classification SCRAMBL1	25
2	Confusion matrix used to produce the classification SCRAMBL2	25
3	Confusion matrix used to produce the classification SCRAMBL3	26
4	Comparison of classifications	28
5	Training field classification accuracy	29
6	Hypothetical data for list sample	42
7	Sample list for CONIFER sample	52
8	Sample list for MEADOW sample	55
9	Calculations for new random number lists	58
10	Ground truth area measurements	61
11	Ground truth area measurements	62
12	Experiment 1 sample data	66
13	Sample summary	70
14	Sample summary	73
15	Sample summary	76
16	Sample summary	79
17	Experiment 4 - summary of all samples	85
18	Sample summary	88
19	Pairwise t test between acreage estimates	91

LIST OF FIGURES

Figure		Page
1	Map of Colorado showing location of study area	8
2	Classification procedure	11
3	Scene showing deciduous (foreground) and conifer (background) classes	13
4	Wolford Mountain, showing non-vegetated/disturbed (foreground), sage (lower mountain), and conifer (upper mountain) classes	13
5	Scene showing mountain shrub class	15
6	Scene showing meadow class	15
7	Scene showing meadow (foreground), mountain grass (light area of background), and sage (background) classes	16
8	Scene showing sage class	16
9	Kremmling, Colorado, showing non-vegetated/disturbed class	17
10	Ellipses of concentration	19
11	C.V. vs. sample unit size	38
12	C.V. vs. sample unit size - CONIFER	49
13	C.V. vs. sample unit size - MEADOW	49
14	C.V. vs. sample unit shape - CONIFER	51
15	90% confidence intervals for acreage estimates - Experiment 1	72
16	90% confidence intervals for acreage estimates - Experiment 2	75
17	90% confidence intervals for acreage estimates - Experiment 3	77

Figure		Page
18	90% confidence intervals for acreage estimates - Experiment 4	82
19	90% confidence intervals for acreage estimates - Experiment 4	86
20	90% confidence intervals for acreage estimates - Experiment 5	89

Plate		Page
1	Map of study area	200
2	Land cover classification map of study area.	201

1.0 INTRODUCTION

The need for assessing the status of the nation's renewable resources has been established by the United States government under Public Law 93-378, the Forest and Rangeland Renewable Resources Planning Act of 1974. A working knowledge of the type, amount, location, and condition of our renewable resources requires rapid inventory techniques which can be frequently updated and compiled with other data. Orbital, aerial, and ground based remote sensing techniques provide a framework for meeting this requirement.

The Middle Park basin and surrounding mountainous areas of north-central Colorado contain land that is important as wildlife habitat, rangeland, and timberland. This area thus provides a specific development and evaluation test site for rapid, wide area remote sensing techniques.

Several remote sensing tools and techniques have been developed and proven effective for obtaining the type of information desired about our renewable resources. Digital satellite data is widely used for coverage of broad areas and is directly suited to computer processing. Aerial photography can be obtained in a wide range of scales to provide the coverage and detail required. Computers are relied on heavily to process the raw remote sensing data and produce information that can readily be used by resource managers.

The prime source of digital satellite data used in natural resources management applications is the Landsat satellies. This data can be processed by any of several computerized algorithms to produce land

cover classification maps. These maps show the location of broad vegetation/land cover categories occurring in an area. Although these maps rarely provide the detail necessary to obtain accurate estimates of the amount and condition of forest and rangeland vegetation types, they may be used as a basis for sampling. Multistage and variable probability sampling techniques, which take advantage of the broad coverage of satellite data and the high degree of detail of aerial photography and on the ground measurements, have been shown to provide accurate estimates over broad areas (Langley, et al., 1969; Nichols et al., 1973; Rhode and Johnson, 1976). The use of digital Landsat data for this purpose is the general topic of this study.

1.1 Study Objective

The purpose of this study is to determine the effects of relative Landsat classification accuracy on acreage estimates made using a variable probability sampling scheme. Such a sampling scheme, also called probability proportional to size (PPS) sampling, provides for a greater probability of selecting those sample units that contain a greater amount of the parameter of interest for further measurement.

The method of using this technique in this experiment was to use the Landsat ground cover classification to gain initial estimates of the acreage of a specific ground cover class. Although the Landsat classification is considered to show the distribution of the various ground cover classes well, acreage estimates from the classification are not considered vary accurate due to the coarse resolution (eight acres per resolution element in this case) and problems inherent in the classification system (e.g., misclassification due to topographic effects,

shadowing, etc.). The study area is thus sampled, and for the selected sample units, detailed measurements of the acreage of the ground cover class of interest are made. These measurements are then used to make an acreage estimate for the entire study area. Simple random sampling could be used for this purpose, in which case the Landsat classification would not be used. A variable probability sampling scheme can be used, however, in which the acreage estimates from the Landsat classification are used to weight the selection probability of the sample units. If the acreage estimates for sample units from the Landsat classification and the acreage measurements from the aerial photography are well correlated, the result will be a decrease in the variance of the final estimate by using variable probability sampling (Langley, 1971).

As previously cited, the benefits of using variable probability sampling in multistage sampling are well documented. One problem that apparently has not been confronted is what effect accuracy of the initial classification has on the final estimate(s) obtained from the sample. If the effect is significant, it would be important to create high quality classifications for use in sampling. If the effect is not significant, substantial savings in time and money could be realized in creating a classification where a high degree of accuracy is not important.

The purpose of this thesis will be to prove or disprove the following hypothesis: Relative classification accuracy has a significant effect on final acreage estimates obtained when using Landsat data to obtain selection probabilities in a variable probability sampling scheme.

The test of this hypothesis will consist of obtaining acreage estimates using different ground cover classifications of the same area, and determining if the estimates thus obtained are dependent upon the relative classification accuracy.

1.2 Experimental Procedure

The first step of the procedure was to create a land cover classification for the Middle Park study area (Ranson and Itkowsky, 1977). The study area selected consists of approximately 589 square miles (1525 km^2). It includes the Colorado River valley between Hot Sulphur Springs and Kremmling and the surrounding mountainous areas.

The classification was obtained using digital Landsat data (frame number 5127-16532) taken August 24, 1975. The data was preprocessed using computer programs available at Colorado State University. This consisted of reformatting, rescaling to a scale of 1:62,500, geometrically correcting, and filtering the data.

The supervised classification technique RECOG (Smith et al., 1972) was used to classify the data into eight broad vegetation/land cover categories. A color coded classification map was manually prepared, and the data maintained in digital form on the computer for further use in the experiment.

Four other classifications, of varying accuracy, were obtained for the same area. One of these was created using a different set of class spectral signatures in the RECOG classification procedure. The other three were obtained from the original classification by randomly reclassifying a certain percentage of points. The relative accuracy of

these classifications was determined by computing the accuracy of classification of training and test areas for which ground truth was known.

Five different sampling experiments were conducted using each of the five classifications to obtain sample unit selection probabilities. Two of these experiments consisted of preliminary samples to estimate conifer acreage. One was a preliminary sample to estimate meadow acreage. Based on the results of these experiments, two additional experiments were conducted for conifer acreage estimation.

The optimum size and shape of sample units were determined from the original classification. Specific sample units were selected by means of list sampling. Ground truth measurements of conifer and meadow acreage were obtained from color infrared aerial photography for the selected sample units. These measurements and the corresponding data from the Landsat classifications were used in calculations to produce acreage estimates for the entire study area. It was then determined whether these estimates were dependent upon relative classification accuracy.

The thesis is arranged as follows: Chapter 2 describes the procedures used to obtin the Landsat classifications of the Middle Park study area. Chapter 3 discusses multistage and variable probability sampling and uses an example to illustrate the sampling procedure and formulas. The experimental data, calculations performed, and results are given in Chapter 4. A discussion of the results and the conclusions drawn are presented in Chapter 5.

2.0 PROCEDURE FOR PRODUCING THE LAND COVER CLASSIFICATIONS OF MIDDLE PARK

Five different Landsat classifications of the same area were obtained to test the hypothesis that relative classification accuracy has a significant effect on subsequent variable probability sampling estimates. These five classifications served as different sources of selection probabilities in the sampling procedure. Two of these classifications (ORIGINAL and JOSESIGS) were created by using two different sets of independently derived class spectral signatures. The other three classifications (SCRAMBL1, SCRAMBL2, and SCRAMBL3) were produced by performing a random relabeling of a certain percentage of points in the ORIGINAL classification.

The production of the ORIGINAL classification constituted a major part of this research. An August 24, 1975 digital MSS Landsat scene (frame number 5127-16532) was used for the study. Color infrared aerial photography at scales of 1:65,000 and 1:130,000 (U.S. Forest Service U-2 Mission #76-096 of June 28, 1976) was used in the training procedure. Classification was performed on the Colorado State University CDC 6400 computer using the supervised maximum likelihood program RECOG, described by Smith, et al. (1974), and the unsupervised program ISOCLS (Minter, 1972). The classification procedure is described in detail in a later section.

A modified clustering approach (Fleming, et al., 1975) was used by José Pernia to produce the class spectral signatures used to generate the JOSESIGS classification (Pernia, 1978).

A computer program called SCRMBL (Appendix D) was written to generate the SCRAMBL1, SCRAMBL2, and SCRAMBL3 classifications from the ORIGINAL classification map.

2.1 Study Site Description

The study site used for the experiment described in this paper, consists of the Middle Park basin and nearby mountains of north-central Colorado, and covers and area of 589 $mi.^2$ (1525 km^2) (Figure 1 and Plate 1). The area is bounded by 40°08'N latitude to the north, 39° 45'N latitude to the south, 106°03' W longitude to the east, and 106° 28' W longitude to the west. It includes portions of Grand, Summit, and Eagle counties.

The prominent cultural features in the area are the towns of Kremmling and Hot Sulphur Springs, and the Williams Fork and Green Mountain Reservoirs. Major rivers and streams include the Colorado River, Blue River, Williams Fork, Big Muddy Creek, and Troublesome Creek. Portions of the Gore Range and the Williams Fork Mountains occur in the southern half of the study area. Wolford Mountain is in the northwestern part of the study area. Elevations range from 7200 ft. (2195 m) at the bottom of Gore Canyon where the Colorado River exits the study area on the western boundary to 13,534 ft. (4125 m) at the summit of Mt.Powell at the extreme southern edge of the area.

Important vegetation types found in the Middle Park area include several species of shrubs and forbs, such as sagebrush (Artemisia spp.), rabbitbrush (Chrysothamnus nauseosus (Pall.) Britt.), snowberry (Symphoricarpos spp.), serviceberry (Amelanchier alnifolia Nutt.),

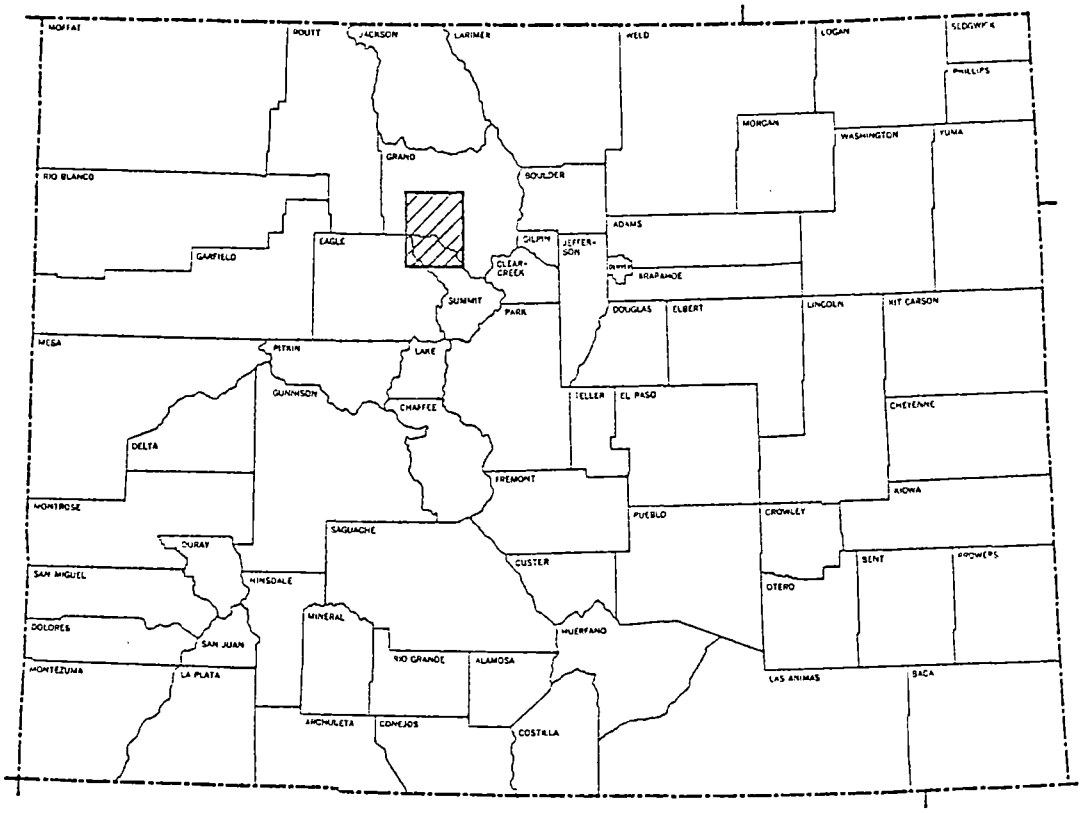


Figure 1. Map of Colorado showing location of study area.

greasewood (Sarcobatus vermiculatus (Hook) Torr.), bitterbrush (Purshia tridentata (Pursh.) DC), Rocky Mountain juniper (Juniperus scopulorum Sarg.), and goldenweed (Haplopapus spp.). Grasses include bluegrass (Pao spp.), wheatgrass (Agropyron spp.), fescue (Festuca spp.), and Indian ricegrass (Oryzopsis hymenoides (R.&S.) Ricker). Sedge (Carex spp.) and rush (Juncus spp.) are also found in meadow areas. The primary forest types of the mountains surrounding the Middle Park basin are quaking aspen (Populus tremuloides Michx.), lodgepole pine (Pinus contorta Dougl. var. latifolia Engelm.), Engelmann spruce (Picea englemanii (Parry) Engelm.), and subalpine fir (Abies lasiocarpa (Hook) Nutt). Other localized or less abundant tree species include Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco), Colorado blue spruce (Picea pungens Engelm.), and cottonwood (Populus angustifolia James). Willow (Salix spp.) and alder (Alnus spp.), along with cottonwood are the primary riparian vegetation types.

The natural resources of the Middle Park area fulfill a variety of management purposes. Of primary importance is the fact that the basin provides a very good critical winter range for the Rocky Mountain mule deer. The rangeland is also used for cattle grazing. The forested mountainous areas provide timber production and recreation. Mining has recently become of importance and there is a large molybdenum mine, the Henderson Mine, along the Williams Fork.

2.2 The ORIGINAL Classification

The computer-assisted classification procedure used to obtain the ORIGINAL classification consists of preprocessing the data, selecting the categories for the classification scheme, obtaining a set of

representative statistics for each category, classifying the Landsat MSS data, and evaluating the performance. This sequence is outlined as a flow diagram in Figure 2.

2.2.1 Data Preprocessing

The first step was to convert the raw Landsat data to the format used by the computer programs at Colorado State University. A program was run to geometrically correct the data, including conversion to a scale of 1:62,500 (Anuta, 1973). This is the scale at which the map is displayed on a line printer printing ten columns per inch and eight lines per inch. The scale of 1:62,500 was chosen because that is the scale of the existing topographic maps for the area. It also produces a map of convenient size and cuts down on the amount of data that need be processed (raw Landsat data is displayed at a scale of approximately 1:24,000). The data values were then smoothed by employing an algorithm developed by Maxwell (1976). This approach calculates a new value for each pixel by taking a weighted average of the eight surrounding points. The effect of this smoothing is to partially correct for atmospheric and sensor noise.

2.2.2 Class Determination/Description

The two main considerations in determining the ground cover classes to be used in the classification were: 1) what types of ground cover exist in the Middle Park study area; and 2) what differences in ground cover can be discriminated using the aerial photography and Landsat data?

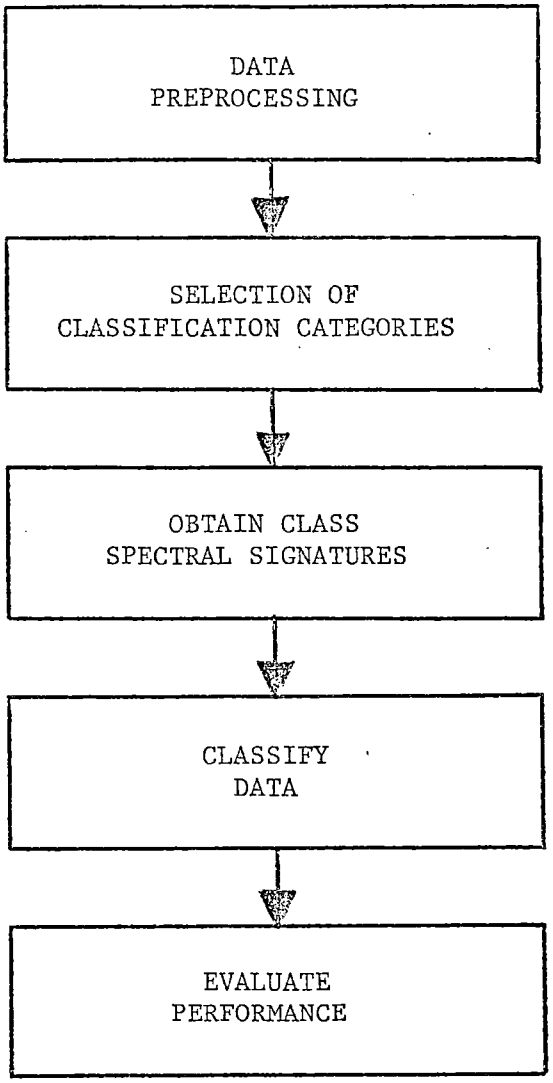


Figure 2. Classification procedure.

As an initial guide, the system of classes used by the Colorado State Forest Service for ground cover mapping was used. Of the 21 classes in the C.S.F.S. system, 13 were found to exist in Middle Park in stands of a size measurable by remote devices. This was determined in a pre-classification reconnaissance trip to the study area. Based on the second consideration, some classes were combined resulting in the eight classes used in the final classification. The classes selected for analysis were:

1. Water - lakes, reservoirs, and rivers. Williams Fork Reservoir and Green Mountain Reservoir are the two prominent examples of this class in the study area.
2. Coniferous forest - vegetation cover on mountain slopes consisting primarily of pure and mixed stands of lodgepole pine (Pinus contorta) and Engelmann spruce/subapline fir (Picea engelmannii/Abies lasiocarpa). Also included in this class could be stands of mixed conifer and aspen and dense stands of pinon-juniper. Good examples of this class are shown in Figures 3 and 4.
3. Deciduous forest - vegetation cover on mountain slopes and along rivers consisting of aspen and cottonwoods (Populus spp.). Dense stands of willows (Salix spp.) might also be included in this class. An example of this class can be seen in Figure 3.
4. Mountain shrub - vegetation cover consisting of deciduous shrub species, e.g., serviceberry (Amelanchier alnifolia), snowberry (Symphoricarpos spp.), sage (Artemisia spp.) and

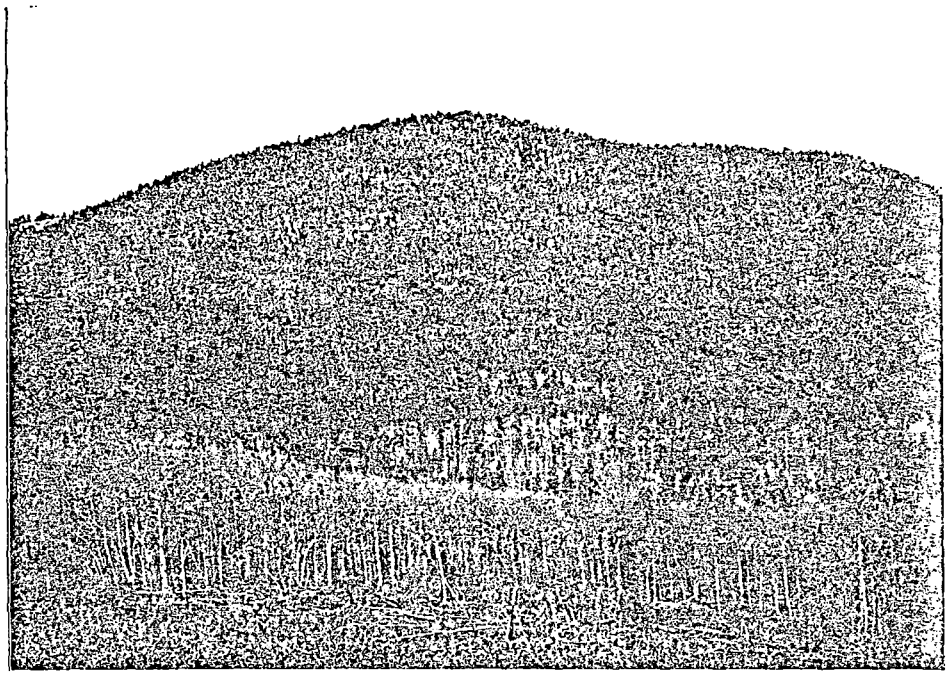


Figure 3. Scene showing deciduous (foreground) and conifer (background) classes.

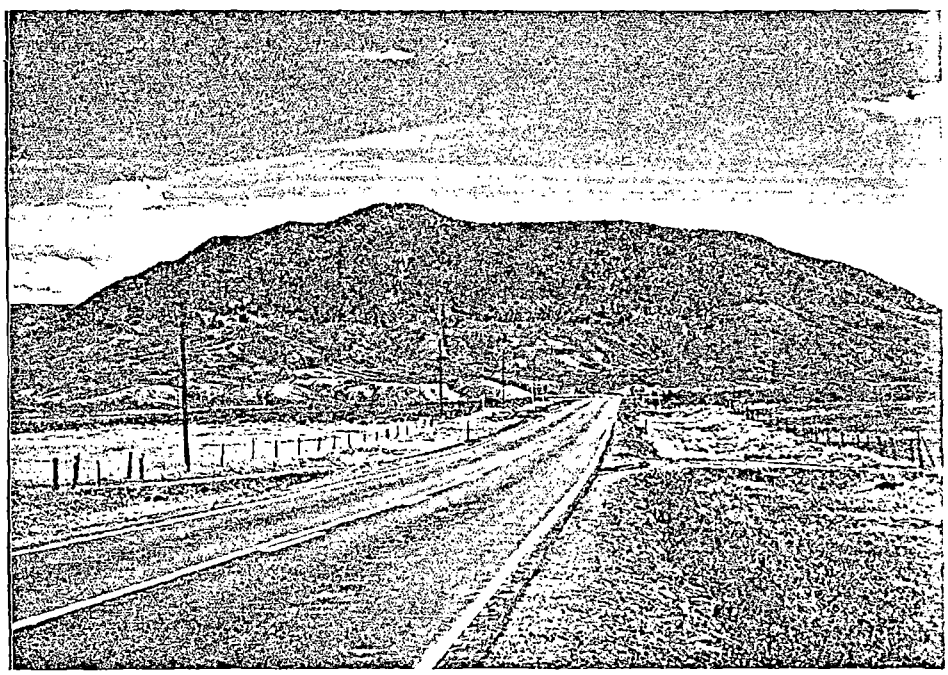


Figure 4. Wolford Mountain, showing non-vegetated/disturbed (foreground), sage (lower mountain), and conifer (upper mountain) classes.

small aspen clumps (Populus tremuloides). Open and mixed stands of juniper (Juniperus spp.) would also be an example of this class. Figure 5 shows an actual training field for this class used for the computerized classification.

5. Meadow - vegetation cover of hayfields, pastures, and other wet fields along streams, irrigated upland fields, and lush mountain meadows, consisting mainly of grass (Carex spp., Juncus spp., etc.). Examples are shown in Figures 6, and 7.
6. Mountain grass - vegetation cover on non-irrigated upland fields consisting mainly of grass (Pao spp., Festuca spp., etc.) and forbs (e.g., penstemon). It is found largely on ridge tops as seen in Figure 7.
7. Sage - vegetation cover consisting mainly of sage (Artemisia spp.) intermixed with Agropyron spp., Poa spp., and Festuca spp. Figures 4, 7, and 8 show this class.
8. Non-vegetated/disturbed - bare rock and soil, large buildings, paved areas and other areas without vegetation. Also included are areas of sparse vegetation (sage or grass) consisting mainly of rock, soil, or dead vegetation. Examples are given in Figures 4 and 9.

2.2.3 Signature Extraction

A line printer graymap of the data was obtained for each Landsat channel. These graymaps display one channel of Landsat data as gray shades determined by the intensity of measured radiation.

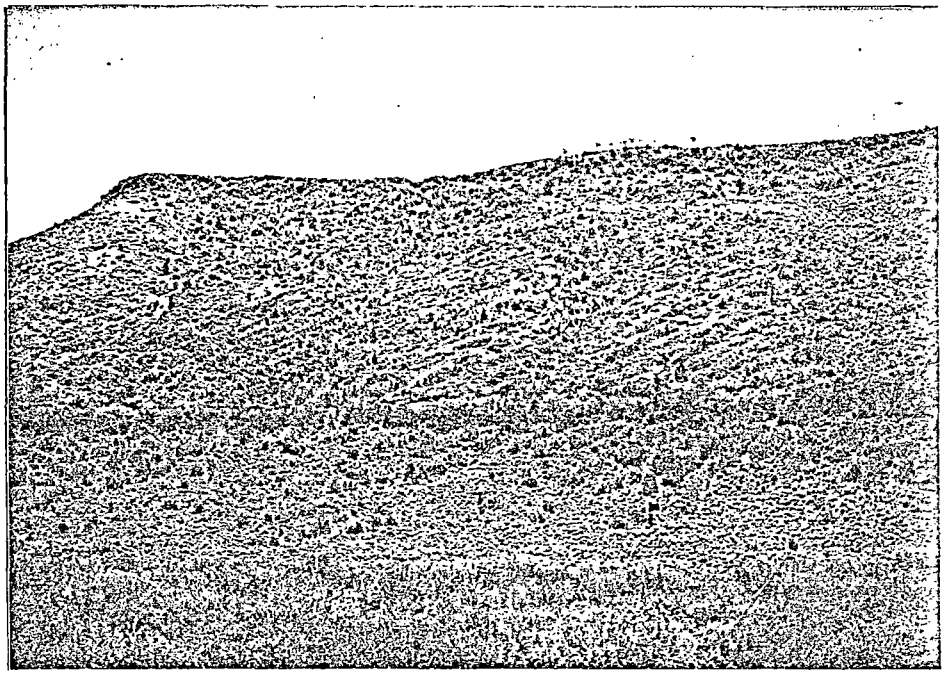


Figure 5. Scene showing mountain shrub class.

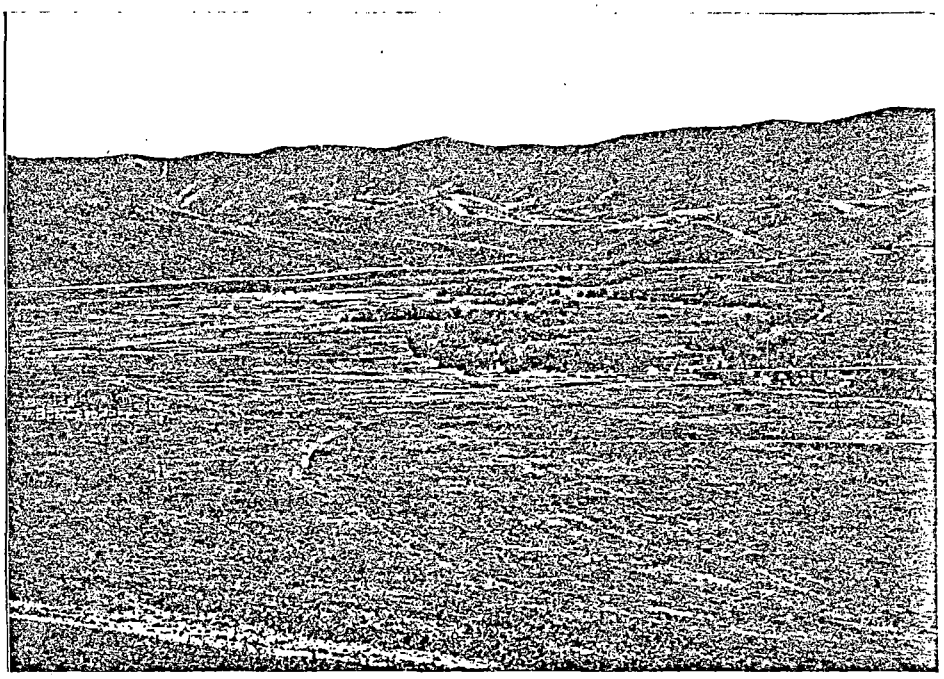


Figure 6. Scene showing meadow class.

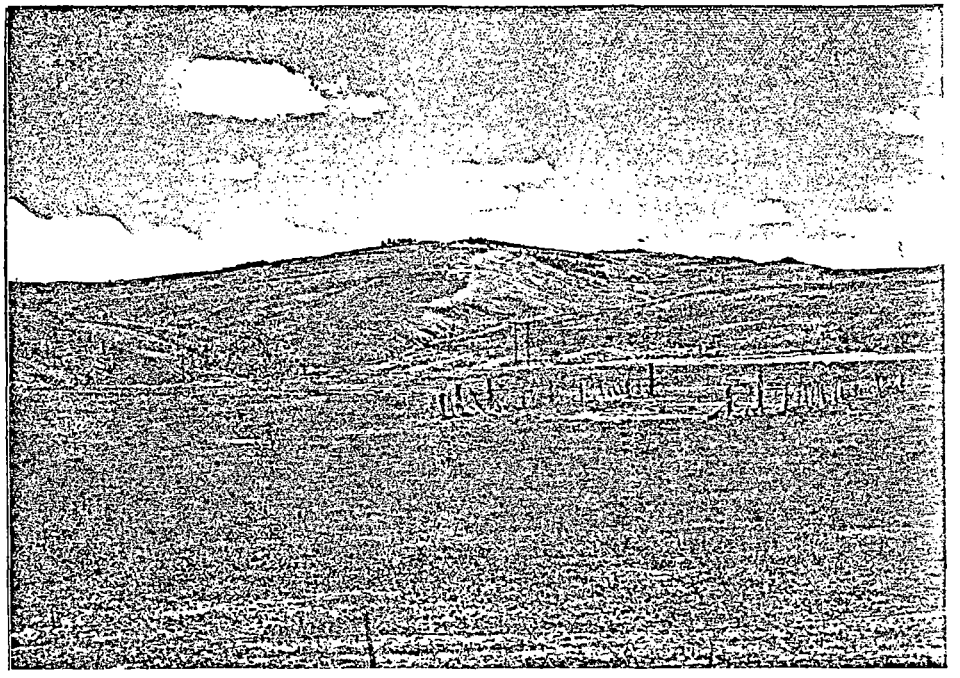


Figure 7. Scene showing meadow (foreground), mountain grass (light area of background), and sage (background) classes.

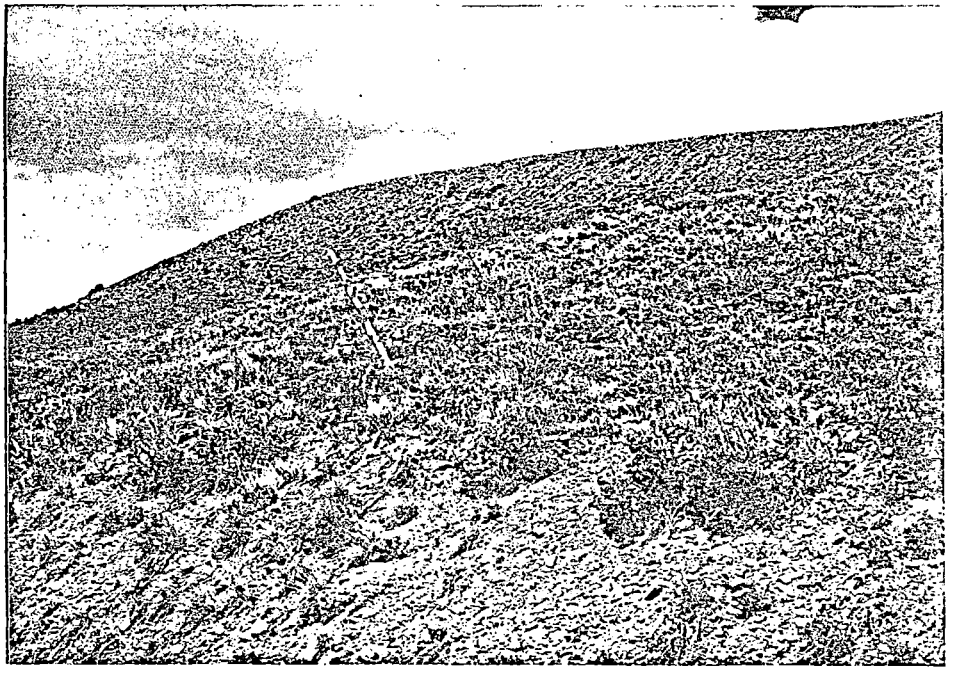


Figure 8. Scene showing sage class.

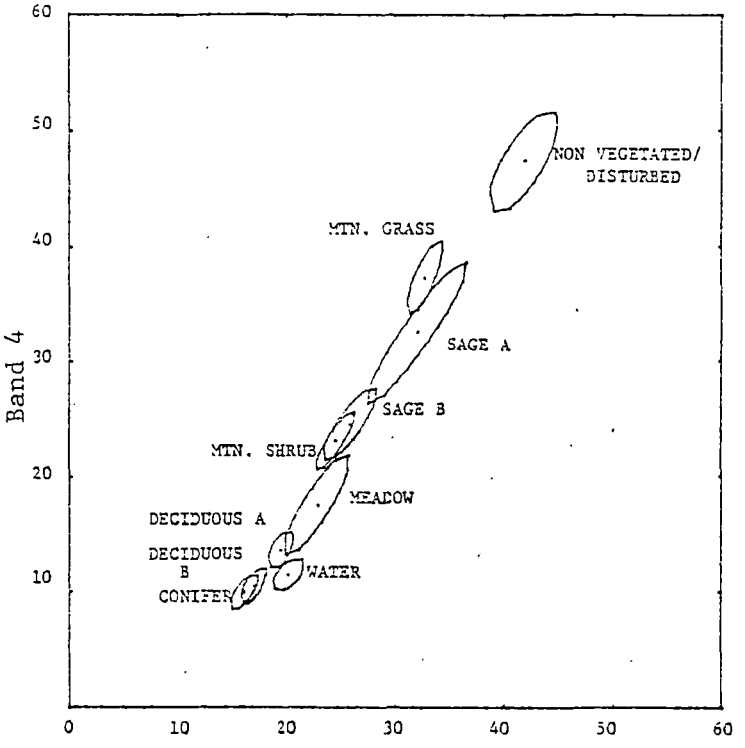
Figure 9. Kremmling, Colorado, showing non-vegetated/disturbed class.

In using the supervised approach, training fields thought to be representative of each of the eight land cover classes were delineated on the graymap. The aerial photography was used to locate areas large enough and homogeneous enough to be used as training fields. An unsupervised clustering approach (Minter, 1972) was used to obtain class signatures where adequate training fields (areas of sufficient size and homogeneity) could not be found. This was the case with the non-vegetated/disturbed class, which is inherently heterogeneous.

Bimodal clusters for the classes deciduous and sage were apparent and these clusters were split into two subclasses. The subclasses were combined after final classification for the final display.

The spectral signature statistics were evaluated by analyzing the relationship of the class data clusters to one another in two-band combinations. A useful method for accomplishing this is to protray the data as an ellipse of concentration with the slope determined by the principal components (eigenvectors) and the axes representing the eigenvalues of the two-band variance-covariance matrix (Davis, 1973). The resulting ellipse shows the data cluster in two-space and includes 66.7% of the class data for those spectral bands.

When the ellipses of concentration for each class are compared for any two-band combination, the spectral separability of the classes can be inferred. Figure 10 shows a series of ellipses of concentration plots for the class statistics in all possible two-band combinations. Upon inspection of these plots, several characteristics of the class spectral signatures are readily apparent. As expected, data for bands 4 vs. 5 and 6 vs. 7 are highly correlated. Data clusters for water,



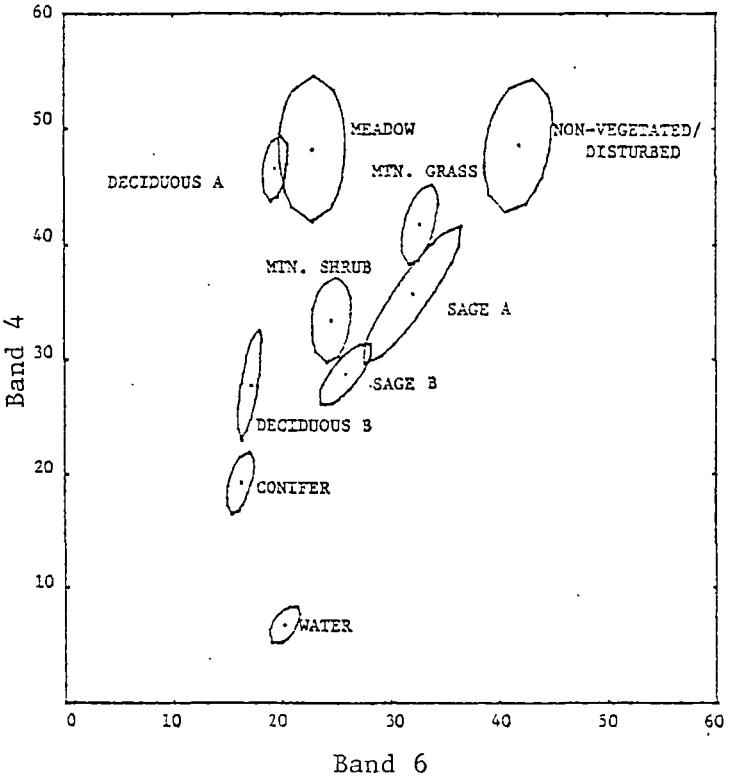
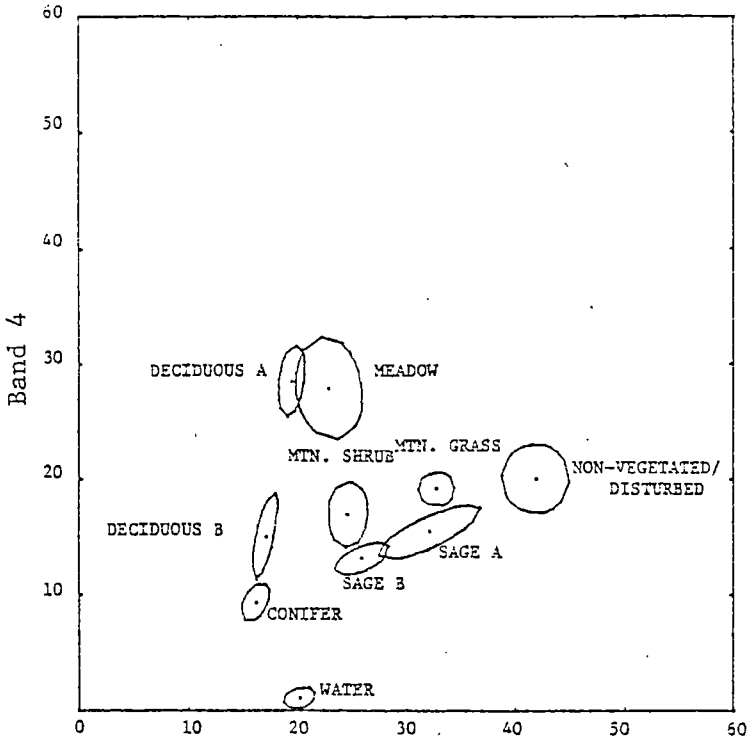


Figure 10. Ellipses of concentration.



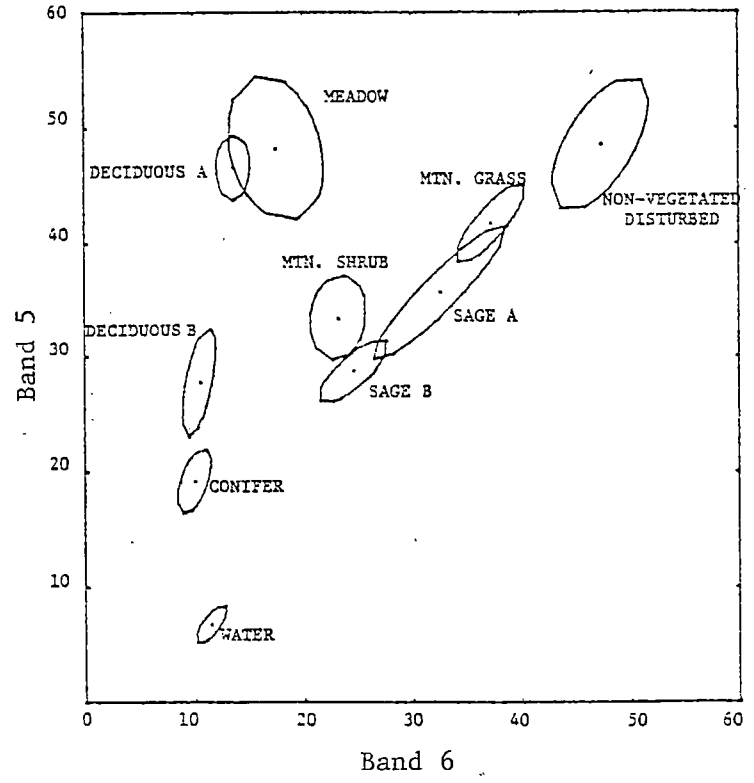
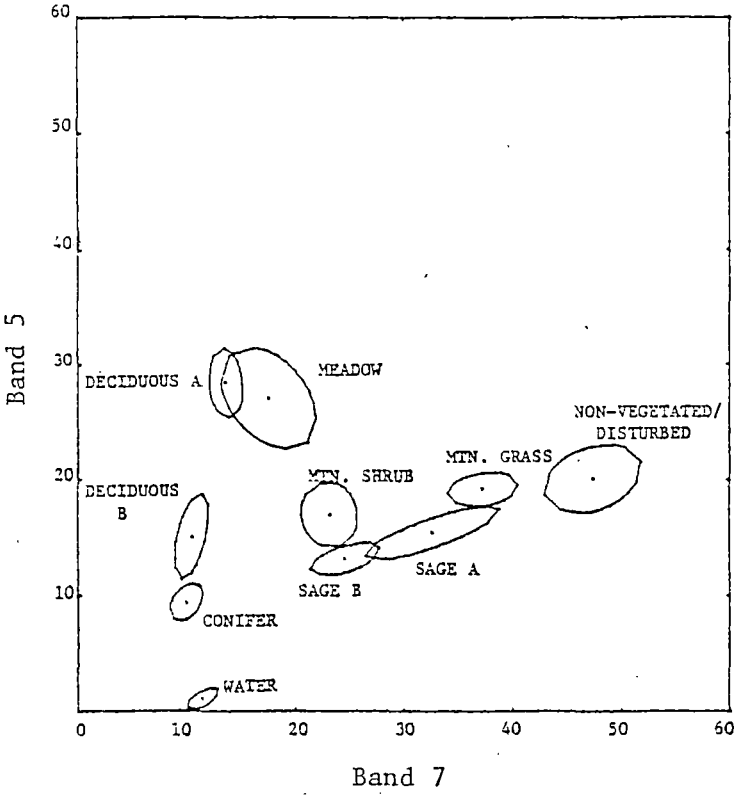


Figure 10. (Continued)



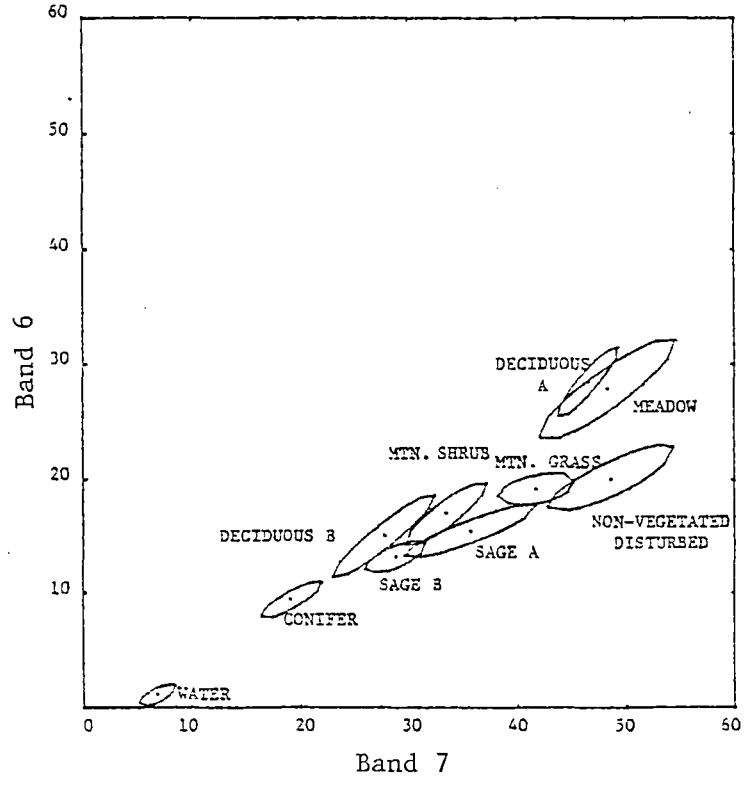


Figure 10. (Continued)

non-vegetated/disturbed, mountain grass and sage A are essentially spectrally distinct for all combinations. Sage B, deciduous B and conifer are spectrally distinct in all combinations but bands 4 vs. 5, meadow and deciduous A, have some overlap of data clusters in all band combinations. Due to the coarse resolution of the 1:62,500 scale data and the difficulties encountered in locating other representative training sets for the deciduous A class, the signature set was left intact for the initial classification of the Landsat MSS data.

2.2.4 Classification and Display

Once the signature extraction/evaluation has been performed the final classification and display are generated. Since analysis of the final display provides perhaps the best method of signature evaluation, the process of signature extraction, evaluation, classification, and display proved to be an iterative one. It did require, however, considerable human and computer time to perform one iteration of this process, so that every effort was made to obtain good class signatures before the final classification and display were performed.

The form of display used is a graymap similar to that produced for display of the original channels. Eight gray levels are used, each corresponding to a land cover class. Each pixel, or printed character, on the classification map represents an area of spproximately 8 acres (20 hectares). Land cover classes were delineated by drawing lines around groups of pixels of the same class. An arbitrary rule of requiring at least four contiguous pixels of the same class to delineate the area was followed. Thus, the minimum area delineated was about 32 acres

(80 hectares). Based on this delineation, an acetate overlay was prepared with each of the eight land cover classes color-coded. This overlay is meant to be used in conjunction with the topographic maps of the area. A reduced representation of the classification is shown in Plate 2.

2.3 The JOSESIGS Classification

A second independent set of class spectral signatures were obtained from José Pernia. These signatures, consisting of a mean vector and covariance matrix for each class were produced from the same Landsat scene in creating a land cover classification of the Williams Fork watershed (Pernia, 1978). A modified unsupervised technique was used to produce the signatures (Fleming et al., 1975). These signatures were used with the Landsat data used for the original classification, to produce a new land cover classification of the Middle Park study area. This independent classification was called the JOSESIGS classification.

2.4 The SCRAMBL1, SCRAMBL2, and SCRAMBL3 Classifications

Classifications of less relative accuracy than the original were obtained by running the original classified data file through a computer program that randomly reclassifies a certain percentage of data points. The program, SCRAMBL, given in Appendix D, was used. The theory behind this program is that if, for example, 85% of the deciduous forest training field data points were correctly classified and 15% were incorrectly classified as mountain shrub, then by randomly reclassifying 15% of the

deciduous points as shrub, a classification will be obtained that is different than the original without a serious degradation of quality.

The exact mechanics of the program are best explained through an example. Input to the program consists of a file of classified data and a confusion matrix representing classification accuracy. Suppose that the confusion matrix of Table 1 is used. This is the ORIGINAL classification training field classification accuracy matrix (Table 20, Appendix A) with arbitrary values added for the non-vegetated/disturbed class. The program looks at each data point in the original classification and, using the information from the confusion matrix row corresponding to the class of a particular point, draws a random number to determine what that point will be reclassified as. For instance, for a point classified as deciduous, the confusion matrix row shows that 92.5% of the deciduous training field points were correctly classified, 3.0% were classified as conifer, and 4.5% were classified as meadow. A random number, x , is drawn from a uniform distribution such that $0 \leq x < 1$. If $x < .925$, the point is reclassified as deciduous (i.e., left the same). If $.925 \leq x < .955$ $(.925 + .030 = .955)$ the point is reclassified as conifer. If $.955 \leq x < 1.0$ $(.925 + .030 + .045 = 1.0)$ the point is reclassified as meadow. If a conifer point is being examined and a random number, x , is drawn, if $0 \leq x < .061$ the point is reclassified as deciduous. If $.061 \leq x < .992$ the point is reclassified as conifer. If $.992 \leq x < 1.0$ the point is reclassified as sage.

The output file of reclassified data can be controlled to a degree (limited by the randomness of the procedure) by the confusion matrix

Table 1. Confusion matrix used to produce the classification SCRAMBL1.

Class	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	.986			.014				
Conifer		.955					.045	
Mt. Grass			.944		.056			
Meadow	.189			.732	.079			
Mt. Shrub					.833		.167	
Water						1.000		
Sage			.011		.021		.968	
Disturbed			.04	.01			.03	.92

Table 2. Confusion matrix used to produce the classification SCRAMBL2.

Class	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	.84	.04		.08	.02		.02	
Conifer	.05	.81				.01	.13	
Mt. Grass			.81		.10		.05	.04
Meadow	.21			.62	.14		.03	
Mt. Shrub	.04				.71		.25	
Water		.04				.96		
Sage	.01		.07		.10		.82	
Disturbed			.11	.03			.08	.78

Table 3. Confusion matrix used to produce the classification SCRAMBL3.

Class	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	.84	.10		.04	.01		.01	
Conifer	.01	.98					.01	
Mt. Grass			.81	.05	.07		.04	.03
Meadow	.04	.05		.90	.01			
Mt. Shrub	.04			.05	.71		.20	
Water		.04				.96		
Sage	.01		.06	.04	.07		.82	
Disturbed			.07	.10			.05	.78

input to the program. The first reclassification (SCRAMBL1) was obtained using the confusion matrix of Table 1. For SCRAMBL2 this confusion matrix was modified to produce a smaller number of conifer and meadow points in the reclassified file (Table 2). This was accomplished by multiplying the diagonal elements of the ORIGINAL confusion matrix by 85% and arbitrarily distributing the rest (so that the sum of a confusion matrix row totals 100%) over classes with similar spectral signatures, as determined from the plots of the ellipses of concentration (Figure 10). SCRAMBL3 was produced by inputting an alteration of the original confusion matrix modified so as to produce a larger number of conifer and meadow pixels in the reclassified data file (Table 3). A chart showing the number of points classified as conifer and meadow in each classification, and the percent change from the original classification is shown in Table 4.

It is believed that the classification data files thus produced maintain the same general distribution of vegetation patterns as the original classification, but produce changes in the classification (and thus in the resulting selection probabilities) that could have resulted by having another person create a classification of the same area (such as JOSESIGS).

2.5 Relative Classification Performance

A quantitative estimation of class separability for the ORIGINAL classification was made by constructing a confusion matrix to show how the training fields were classified. Confusion matrices are shown in Table 5. The non-vegetated/disturbed class is not shown in the matrices

Table 4. Comparison of classifications.

CONIFER

Classification	Number of CONIFER Points	% of Total Area	Change in % from ORIGINAL
ORIGINAL	5774	14.4	-
JOSESIGS	8708	21.8	50.8
SCRAMBL1	5509	13.8	-4.6
SCRAMBL2	5032	12.6	-12.9
SCRAMBL3	6637	16.6	14.9

MEADOW

Classification	Number of MEADOW Points	% of Total Area	Change in % from ORIGINAL
ORIGINAL	4526	11.3	-
JOSESIGS	2216	5.5	-51.0
SCRAMBL1	3452	8.6	-23.7
SCRAMBL2	3450	8.6	-23.8
SCRAMBL3	5510	13.8	21.7

Table 5. Training field classification accuracy.

Points Classified As:

Training Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	62	2		3				
Conifer	8	121					1	
Mt. Grass			17				1	
Meadow	28		1	93	6			
Mt. Shrub					10		2	
Water						129	1	
Sage			7		4		169	6

Percentage Classified As:

Training Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	92.5	3.0		4.5				
Conifer	6.1	93.1					0.8	
Mt. Grass			94.4				5.6	
Meadow	21.9		0.8	72.7	4.7			
Mt. Shrub					83.3		16.7	
Water						99.2	0.8	
Sage			3.8		2.2		90.9	3.2

Percent of points correctly classified = 89.6

because, as previously explained, training fields were not used to obtain the class signature.

It can be seen that classification of the training fields was better than 90% correct for all classes except mountain shrub and meadow. In the case of mountain shrub, this is because of the small size of the training field (only two points were misclassified resulting in an accuracy of 83.3%). In the case of meadow, 18.9% of the points were misclassified as deciduous and 7.9% as mountain shrub. The confusion with deciduous can be explained by looking at the ellipses of concentration (Figure 10). In each plot the ellipse for the meadow class is relatively large and overlaps the ellipse for the deciduous A class.

A qualitative analysis of the total scene classification was made by using the aerial photographs and a familiarity with the study area to test the classification results.

The observation was made that the classification map shows a definite gradation from sage to mountain shrub to deciduous forest to coniferous forest. This gradation is obviously related to elevation. In actuality there is no sharp demarcation between the areas occupied by these vegetation classes. The sparse sage of drier lower elevations becomes denser and mixed with other shrub species at higher elevations. At higher elevations aspen gradually becomes more prominent. At still higher elevations the aspen mixes with pine, spruce, and fir and eventually gives way to pure stands of spruce/fir.

The relative accuracy of the four new classifications was determined by obtaining the classification accuracy of test fields. The

test fields used were the same areas used as training fields for the ORIGINAL classification. The confusion matrices showing the test field classification accuracy for the four new classifications are shown in Tables 21 through 24 in Appendix A. More rigorous tests of absolute classification accuracy were not performed because these tests are generally dependent upon very good registration between the Landsat classification map and the ground truth (topographic) map. The registration between the two maps in this experiment is considered to be accurate, at best, to one eight acre pixel. Although this is good enough to satisfy the sampling assumption of good correlation between the Landsat and ground truth data, it could cause large errors in the results of most techniques used to estimate classification accuracy.

It is believed that the method of estimating classification accuracy used here is sufficient to show relative differences between classifications. The purpose of the study is to show whether or not using classifications of varying accuracy produces significantly different sample estimates. No attempt was made to correlate these estimates with absolute classification accuracy.

3.0 MULTISTAGE AND VARIABILITY PROBABILITY SAMPLING

Detailed measurements of natural resources parameters can be made in a number of ways, making use of the wide variety of imagery available. Landsat imagery, for instance, provides the wide coverage necessary to make measurements over broad areas at a reasonable cost. Its low degree of accuracy due to its coarse resolution and classification errors, however, may preclude its use. Medium to large scale aerial photography and on the ground measurements can provide a very high degree of accuracy, but are not economically feasible for use over broad areas. Therefore, sampling, making use of the broad coverage of small scale imagery (e.g., Landsat) and the detail of larger scale imagery and on the ground measurements, is considered a more feasible method (Langley, 1969 and 1971).

In forest surveys stratified sampling, double sampling with stratification at the first phase, and double sampling with regression have been employed to make use of the different levels of information. Another technique, developed recently, is multistage sampling with arbitrary probabilities of selection at each stage. In such a technique the sampling probabilities are formulated from the additional information available by virtue of the increasingly finer resolution of remote sensor data at each stage (Langley, 1969 and 1971). Multistage sampling using satellite data or small scale aerial photography in the first stage has been used by a number of researchers and proven cost effective in forest inventories (Nichols, et al., 1973), crop acreage estimation (Rhode and Johnson, 1976), shrub community parameter estimation (Driscoll

and Francis, 1970), forest insect damage estimation (Heller and Wear, 1969), and other applications. Perhaps the greatest benefit realized by using this technique is a reduction in variance of the final estimate when compared to other techniques. The result of this reduced variance is a reduction in the size of the sample needed to produce a given degree of accuracy (Nichols, et al., 1973).

An example of a three stage sample to estimate timber volume on a large tract of land would be as follows: Landsat data is used to determine the selection probabilities of sample units based on estimates of timber acreage in the Landsat classification. The first stage sample, using these selection probabilities, is then conducted using medium scale aerial photography (e.g., at a scale of 1:24,000). More detailed estimates of timber acreage are obtained and this information is used to determine the selection probabilities at the next stage. The next stage of sampling would be conducted using large scale aerial photography (e.g., 1:2000 scale) from which timber volume estimates are made. These volume estimates are then used to determine selection probabilities for the third stage of sampling. This stage would involve making on the ground measurements of tree volume using a dendrometer. These measurements are then projected back through the sampling formula to obtain an estimate of timber volume for the whole area.

Multistage sampling is able to provide estimates with small variance through the use of variable probability sampling, also called probability proportional to size (PPS) sampling. This technique was developed in 1943 by Hansen and Hurwitz (1943). It involves giving those sample units that contain a greater amount of the parameter of interest

a greater probability of being selected. This is accomplished by using known information to determine the selection probabilities of sample units. All that is required to realize a decreased variance using this method (versus other sampling methods), is that the known information be well correlated with the parameter of interest. In the above example it is assumed that timber acreage estimates from Landsat are well correlated with acreage estimates from the medium scale aerial photography, the acreage estimates from the medium scale photography are well correlated with volume estimates from the large scale photography, and that the volume estimates from the large scale photography are well correlated with the on-the-ground volume measurements. As long as this assumption holds true, the variance of the final estimate will be relatively small (Langley, 1971).

The following discussion of variable probability sampling and its advantage over sampling with equal probabilities is taken mainly from the work of Langley (1971). In simple random sampling, an estimate of the total resource quantity on a tract of land is given by:

$$\hat{A} = \frac{N}{n} \sum_{i=1}^{n} y_i$$

where N = the number of units in the population

n = the number of sample units

y_i = the quantity of the resource measured in the i^{th} sample unit.

This may be rewritten as:

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{1/N}$$

where 1/N is the probability, p, of drawing the i^{th} unit in a single draw. Therefore:

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p}$$

where each ratio, y_i/p, estimates the population total. The true variance of this estimator (not the estimated variance) when sampling with replacement is:

$$\text{Var}(\hat{A}) = \frac{1}{n} \sum_{i=1}^{n} p\left(\frac{y_i}{p} - A\right)^2$$

where A is the true resource value. The variance here arises from the variation among the population units, y_i, and is unaffected by the constant probability of selection.

In variable probability sampling, the probabilities of selection are no longer a constant, 1/N, but are defined as:

$$p_i = \frac{x_i}{\sum_{i=1}^{n} x_i}$$

where x_i is a prediction of the resource quantity contained in the i^{th} population unit and $\sum_{i=1}^{n} x_i$ is the cumulative sum of x_i over the entire population. The estimator may now be written as:

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i}$$

with variance

$$\text{Var}(\hat{A}) = \frac{1}{n} \sum_{i=1}^{n} p_i\left(\frac{y_i}{p_i} - A\right)^2$$

Here again, each ratio y_i/p_i estimates the population total. In this case, however, high linear correlation between y_i and p_i causes the ratios y_i/p_i over all units to be similar and close to the true population total, A. Therefore, the variance is lower than with equal probability sampling due to the smaller deviations of the ratios y_i/p_i from A. Perfect correlation between y_i and p_i gives a variance of zero, regardless of sample size. Negative correlation, on the other hand, gives a variance greater than with simple random sampling.

The estimator $\hat{A}$ is unbiased as long as the sample units are physically drawn with probability proportional to the prediction. This is accomplished by using list sampling (Husch, et al., 1972).

As an illustration of the mechanics of performing a variable probability sample a simple one stage example is given. Assume that it is desired to estimate the acreage of coniferous forest in a specific study area. In this example a Landsat classification of the study area is used to obtain selection probabilities of sample units and color infrared aerial photography at a scale of 1:10,000 is used to obtain detailed acreage estimates. The Landsat image contains 100 lines and 100 columns, a total of 10,000 pixels, with each pixel representing approximately 1 acre.

The first step is to determine the size and shape of sample units to use. These are dependent upon the amount of variation resulting from a particular size and shape and practical considerations (Rhode, 1976). The amount of variation is inversely related to sample unit size, which favors using large sample units. A practical consideration would be to keep the sample units small enough so that most of them

would be covered by a single frame of the aerial photography. For the sake of simplicity, only square sample units (i.e., width in pixels to length in pixels) are considered.

With no prior information, the only way to determine the amount of variation that results from a particular size sample unit is by conducting an actual sample (Bowden, 1978). Since this would be highly impractical due to the large number of possible sample unit sizes, another method is used. This is to grid off the Landsat classification into sample units of various sizes, count the number of pixels in each sample unit classified as conifer, and obtain a mean and corresponding standard deviation and coefficient of variation. The equations used are:

$$\text{mean} = \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{3.1}$$

where: x_i = count of conifer pixels in the sample unit

N = number of sample units in the entire study area

$$\text{standard deviation} = s = \sqrt{s^2} \tag{3.2}$$

where:

$$s^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \tag{3.3}$$

$$= \frac{1}{N} \left[\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} x_i \right)^2 \right] \tag{3.4}$$

$$\text{coefficient of variation} = \text{C.V.} = s/\bar{x} \tag{3.5}$$

The coefficient of variation is computed for various size sample units and a plot of C.V. versus sample unit size made (Figure 11). The number of possible sizes is limited by the requirement of having an

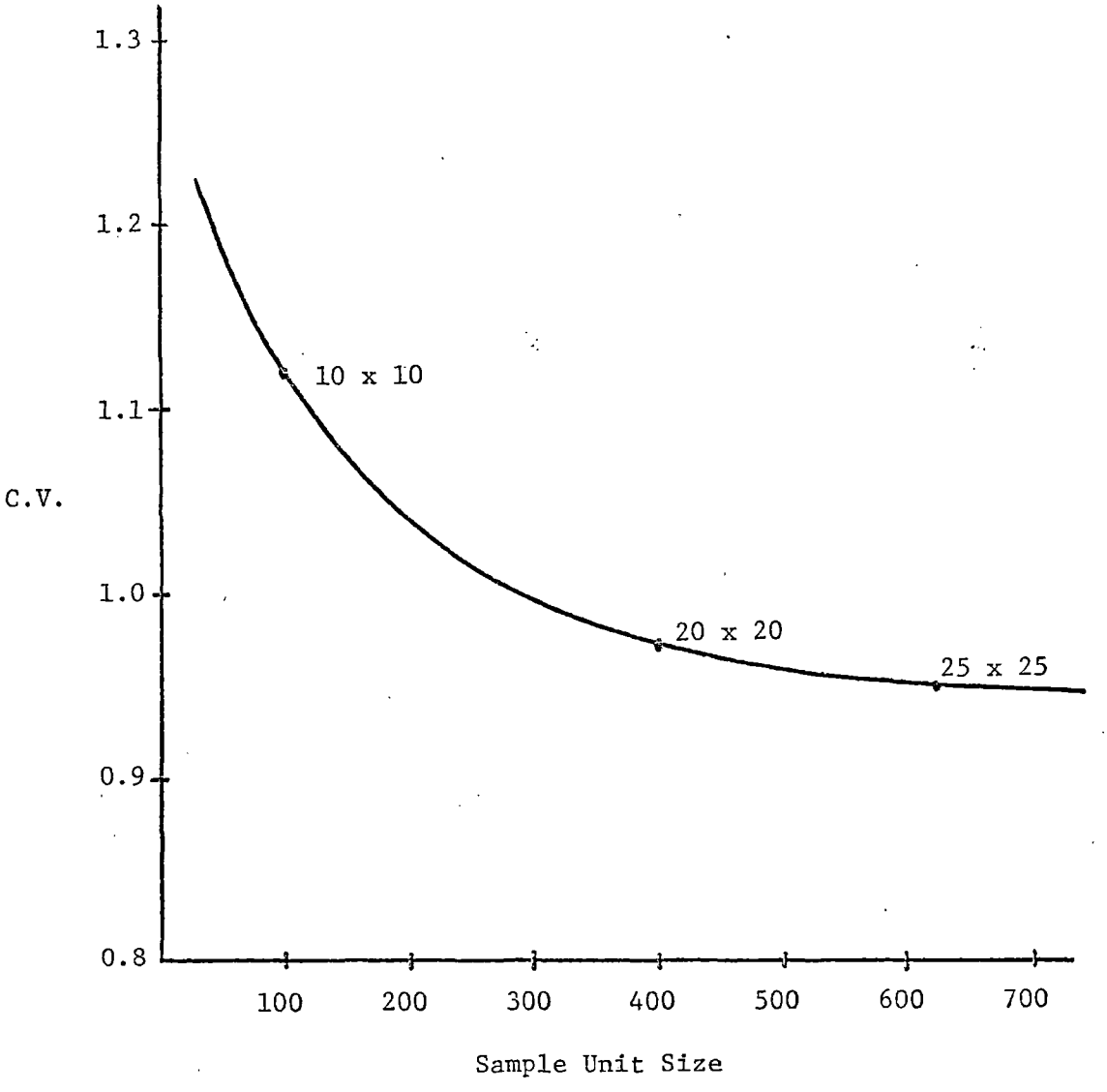


Figure 11. C.V. vs. sample unit size.

integral number of sample units in the study area. The point on the curve where an increase in sample unit size does not produce a significant decrease in C.V. (i.e., where the curve starts to flatten out), within the limits established by practical considerations, is the optimum size sample unit. In this example it was decided to use sample units 20 pixels long by 20 pixels wide.

The purpose of this procedure to determine optimum size sample units is to reduce the variance of the final estimate. The variance (and coefficient of variation) calculated as shown above, however, is not necessarily related to the variance of the final estimate. The only way to determine this latter variance is to conduct actual samples. Since it would be impractical to conduct samples to test the effect of every possible sample unit size and shape, the above procedure was considered to be adequate for this purpose.

A greater determinant of the variance of the final estimate than sample unit size and shape is the sample size (number of sample units included in the sample). This is why determination of optimum size and shape of sample units was not considered to be a critical factor.

After the size and shape of sample units has been determined, the sample can be conducted. The number of sample units to select for the actual sample is dependent upon the variance of the final estimate and the degree of accuracy desired. Rhode (1976) gives a method for calculating the number of sample units to include in the sample to obtain a specified degree of accuracy. The equation used is:

$$n = \frac{t^2 \ C.V.^2}{A.E.^2}$$

where: n = number of sample units to include in the sample

t = t statistic at n-1 degrees of freedom and given confidence level

C.V. = coefficient of variation

A.E. = allowable error

The coefficient of variation used in this formula is that obtained from the Landsat classification by the procedure described in the previous section. As mentioned, this is not the same as the coefficient of variation of the final estimate, which should be used to determine sample size. Use of the coefficient of variation obtained from the Landsat classification results in gross over-estimation of sample size since this value is usually several times larger than that obtained for the final estimate (on the order of 100% versus 10%). The way to determine the correct value for the coefficient of variation is by actually conducting a sample and obtaining an acreage estimate and corresponding variance and coefficient of variation (Bowden, 1978). The same equation can then be used to determine the sample size needed for the specified accuracy.

It is thus decided to conduct a preliminary sample of five sample units. When this is completed a coefficient of variation and required sample size can be computed. If the necessary sample size is five or less, no further sampling will be necessary. Additional sample units will be selected if it is greater than five.

The first step is to grid off the Landsat classification into sample units and count the number of conifer pixels in each sample unit. A list is drawn up showing: sample unit number, i; count of conifer pixels in sample unit i, x_i; cumulative sum for sample unit i,

$\sum_{j=1}^{i} x_j$ (i.e. the sum of x_i for all sample units up to and including sample unit i); and probability of selection of sample unit i, p_i ($p_i = x_i / \sum_{i=1}^{N} x_i$, where $\sum_{i=1}^{N} x_i$ is the cumulative sum of x_i for all sample units). This list is shown in Table 6.

To determine the specific sample units to be included in the sample five random numbers between 0 and $\sum_{i=1}^{N} x_i$ (equal to 1000 in this example) are drawn. These random numbers are compared to the cumulative sum list ($\sum_{j=1}^{i} x_j$) of Table 6, and the sample units whose interval includes one of the random numbers are selected for the sample. Sampling is done with replacement, so it is possible to select the same sample unit more than once.

Assume that the five random numbers drawn are 532, 904, 762, 201, and 716. Sample unit 9 is included because the random number 532 is greater than 419, the cumulative sum for the sample unit 8, and less than or equal to 595, the cumulative sum for sample unit 9. Likewise, for the random number 904 sample unit 20 is included, for 762 sample unit 15 is included, for 201 sample unit 6 is included and for 716 sample unit 14 is included.

The next step is to obtain detailed acreage estimates for the selected sample units using the aerial photography. In practice this might involve delineating the selected sample units on a topographic base map, determining which aerial photography frame(s) cover the sample unit, delineating the areas of coniferous forest on that frame(s), transferring this information to the topographic base map using a zoom transfer scope, and measuring the area of coniferous forest on the base map with a planimeter or dot grid overlay.

Table 6. Hypothetical data for list sample.

Sample Unit	x_i	$\sum_{j=1}^{i} x_j$	p_i
1	7	7	7/1000 = .007
2	15	22	15/1000 = .015
3	14	36	14/1000 = .014
4	38	74	38/1000 = .038
5	57	131	57/1000 = .057
6	91	222	91/1000 = .091
7	85	307	85/1000 = .085
8	112	419	112/1000 = .112
9	176	595	176/1000 = .176
10	21	616	21/1000 = .021
11	36	652	36/1000 = .036
12	22	674	22/1000 = .022
13	18	692	18/1000 = .018
14	43	735	43/1000 = .043
15	29	764	29/1000 = .029
16	17	781	17/1000 = .017
17	19	800	19/1000 = .019
18	32	832	32/1000 = .032
19	20	852	20/1000 = .020
20	64	916	64/1000 = .064
21	2	918	2/1000 = .002
22	33	951	33/1000 = .033
23	18	969	18/1000 = .018
24	22	991	22/1000 = .022
25	9	1000	9/1000 = .009

When the acreage measurements have been made for the selected sample units, an estimate of coniferous forest acreage for the entire study area and a sample variance and standard deviation are computed using the following equations:

$$\text{acreage estimate} = \hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i} \tag{3.6}$$

where: n = number of sample units

y_i = acreage measurement for sample unit i

p_i = selection probability of sample unit i (from Table 6)

$$\text{sample variance} = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{y_i}{p_i} - \hat{A}\right)^2 \tag{3.7}$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} \left(\frac{y_i}{p_i}\right)^2 - \frac{1}{n} \left(\sum_{i=1}^{n} \frac{y_i}{p_i} \right)^2 \right] \tag{3.8}$$

$$\text{sample standard deviation} = s = \sqrt{s^2} \tag{3.9}$$

As mentioned previously, the true variance of the estimator, $\hat{A}$, is given by:

$$\text{Var}(\hat{A}) = \frac{1}{n} \sum_{i=1}^{n} p_i \left(\frac{y_i}{p_i} - A\right)^2 \tag{3.10}$$

Since the true acreage, A, is unknown, an estimator of this variance (termed the population variance, where population now refers to a population of estimators $\hat{A}$) must be used. This is given by the equation (Bowden, 1978; Cochran, 1963):

$$\hat{\text{Var}}(\hat{A}) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \left(\frac{y_i}{p_i} - \hat{A}\right)^2 \tag{3.11}$$

This is merely the sample variance divided by the sample size:

$$\hat{\text{Var}}(\hat{A}) = s^2/n \tag{3.12}$$

An estimator of the population standard deviation is given by:

$$\hat{\sigma} = \sqrt{\hat{\text{Var}}(\hat{A})} = s/\sqrt{n} \tag{3.13}$$

This is used to calculate the coefficient of variation:

$$\text{C.V.} = \hat{\sigma}/\hat{A} \tag{3.14}$$

Assume that for the sample of this example the following acreage measurements are obtained: sample unit 6 - y_i = 81 acres; sample unit 9 - y_i = 198 acres; sample unit 14 - y_i = 40 acres; sample unit 15 - y_i = 23 acres; sample unit 20 - y_i = 72 acres. The following calculations are then performed:

$$\hat{A} = \frac{1}{5}\left(\frac{81}{.091} + \frac{198}{.176} + \frac{40}{.043} + \frac{23}{.029} + \frac{72}{.064}\right)$$

$$= 973 \text{ acres}$$

$$s^2 = \frac{1}{4}\left[\left(\frac{81}{.091}\right)^2 + \left(\frac{198}{.176}\right)^2 + \left(\frac{40}{.043}\right)^2 + \left(\frac{23}{.029}\right)^2 + \left(\frac{72}{.064}\right)^2\right.$$

$$\left. -\frac{1}{5}\left(\frac{81}{.091} + \frac{198}{.176} + \frac{40}{.043} + \frac{23}{.029} + \frac{72}{.064}\right)^2\right]$$

$$= 21818$$

$$s = \sqrt{21818} = 148$$

$$\hat{\text{Var}}(\hat{A}) = 21818/5 = 4364$$

$$\hat{\sigma} = \sqrt{4364} = 66$$

$$\text{C.V.} = 66/973 = .068 \text{ or } 6.8\%$$

Thus, the estimated acreage of conifer forest in the entire study area is 973 acres and the coefficient of variation for this estimate is 6.8%.

It is now possible to compute the sample size necessary to achieve a specified degree of accuracy. This is computed by the following equations:

$$n' = \frac{t^2 \ C.V.^2}{A.E.^2} \tag{3.15}$$

where: t = t statistic at n-1 degrees of freedom and specified confidence level

C.V.= coefficient of variation

A.E.= user specified allowable error

Applying a finite population correction factor the resulting required sample size is:

$$n = \frac{n'}{1 + \frac{n'}{N}} \tag{3.16}$$

where N is the number of possible sample units in the population. The value of n computed is rounded up to the nearest whole number.

For the example given assume an allowable error of 10% is specified. The t statistic at 5-1 = 4 degrees of freedom, assuming a 90% confidence level is equal to 2.132. The calculations for required sample size are then:

$$n' = \frac{(2.132)^2 \ (.068)^2}{(.10)^2} = 2.10$$

$$n = \frac{2.10}{1 + \frac{2.10}{25}} = 1.94 \simeq 2$$

Applying the finite population correction factor did not change the required sample size much in this case. For larger values of n', however, it would provide for a reduction in required sample size.

The sample size required for the specified degree of accuracy is equal to two. Since the preliminary sample included five sample units, no further sampling is necessary.

The example given here is a simple one stage sample. The literature contains explanations of how this theory is extended to multistage samples (Langley, 1969 and 1970; Langley, et al., 1969; Nichols, et al., 1973).

4.0 DESCRIPTION OF VARIABLE PROBABILITY SAMPLING EXPERIMENT

Variable probability sampling can be used to obtain estimates of the amount of one or more parameters over a given tract of land. A one stage variable probability sampling scheme is easily implemented for providing an acreage estimate of a ground cover class, as was done in the experiment described here. Multistage schemes are used to provide estimates of more detailed parameters, for example the basal area and volume of a stand of timber (Titus, et al., 1975). A one stage sample was used in this experiment for the sake of simplicity, in the hope that the effects of classification accuracy would be more easily recognized.

For this experiment samples were performed for two land cover classes. Their purpose was to estimate the acreage (area of coverage) of coniferous forest and meadow land in the Middle Park study area. The same Landsat data and procedures were used in each sample, but the samples were performed separately. Samples for two different land cover classes were made in an attempt to note what differences occur, if any, when estimating different parameters. The two ground cover classes were chosen to represent extremes of classification accuracy. Their accuracy, as determined by training field classification accuracy for the original classification, was 95.5% for conifer and 73.2% for meadow.

A subset of the study area was used for these samples. The Landsat classifications of the area consist of 220 columns by 220 lines of data. For the sampling experiments an area represented by the first

200 columns by the first 200 lines of data was used. This corresponds to an area of 487 mi^2 (1260 km^2). Since the study area must be divided into an integral number of sample units, the 200 by 200 subset was used because it presented a greater variety of possible sample unit sizes than the 220 by 220 study area.

4.1 Selection of Sample Unit Size and Shape

The size and shape of sample units is determined by the amount of variation present and practical considerations. The procedure followed to determine size and shape was that described in the previous chapter.

A computer program, PSU (Appendix D), was written to grid off a Landsat classification into sample units of a specified size and shape, make a count of the number of pixels in each sample unit of the class of interest, and compute a mean, standard deviation, and coefficient of variation. Since the number of possible sizes and shapes of sample units is quite large, only square sample units (width in pixels equal to length in pixels) were considered initially.

The program PSU was run a number of times for both conifer and meadow, specifying different size sample units. The resulting coefficients of variation were then plotted versus sample unit size (Figures 12 and 13). The point on this curve where an increase in sample unit size does not produce a significant decrease in coefficient of variation (i.e., where the curve starts to flatten out) is considered to be the optimum size sample unit. Thus, a sample unit size of 20 pixels by 20 pixels was considered optimum for conifer, and 25 pixels by 25 pixels for meadow.

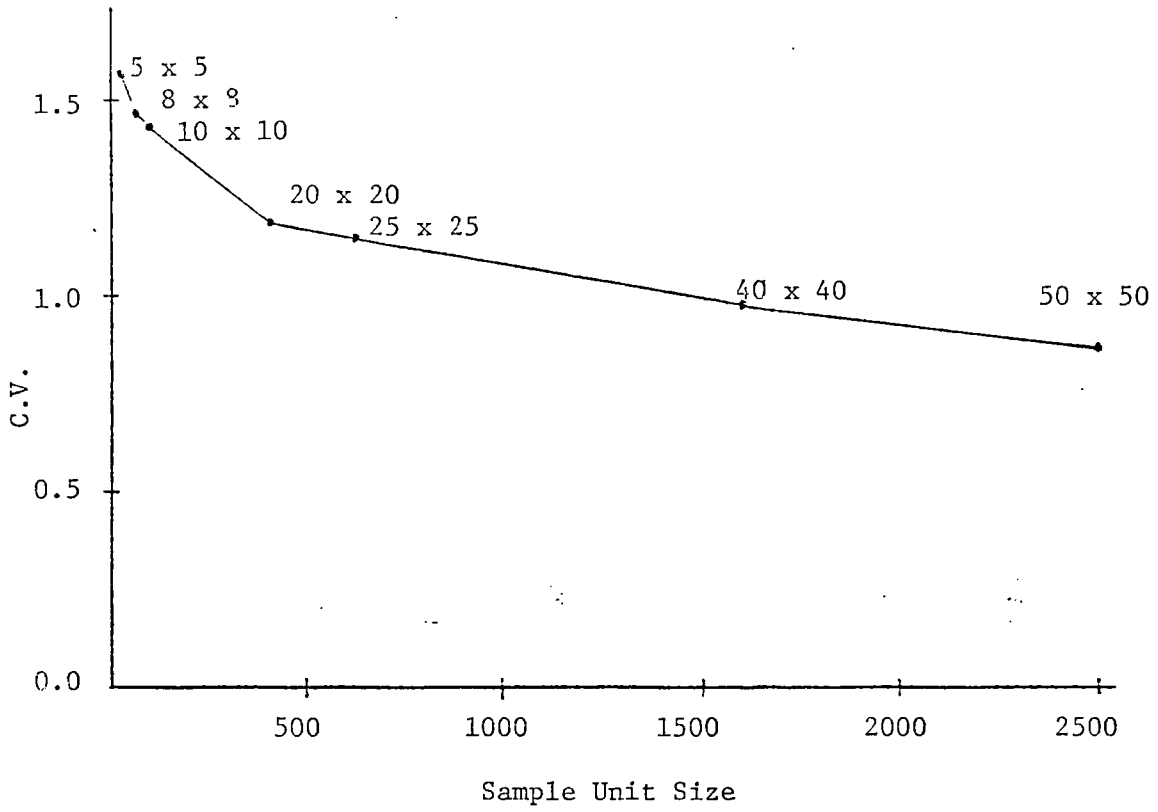


Figure 12. C.V. vs. sample unit size - CONIFER.

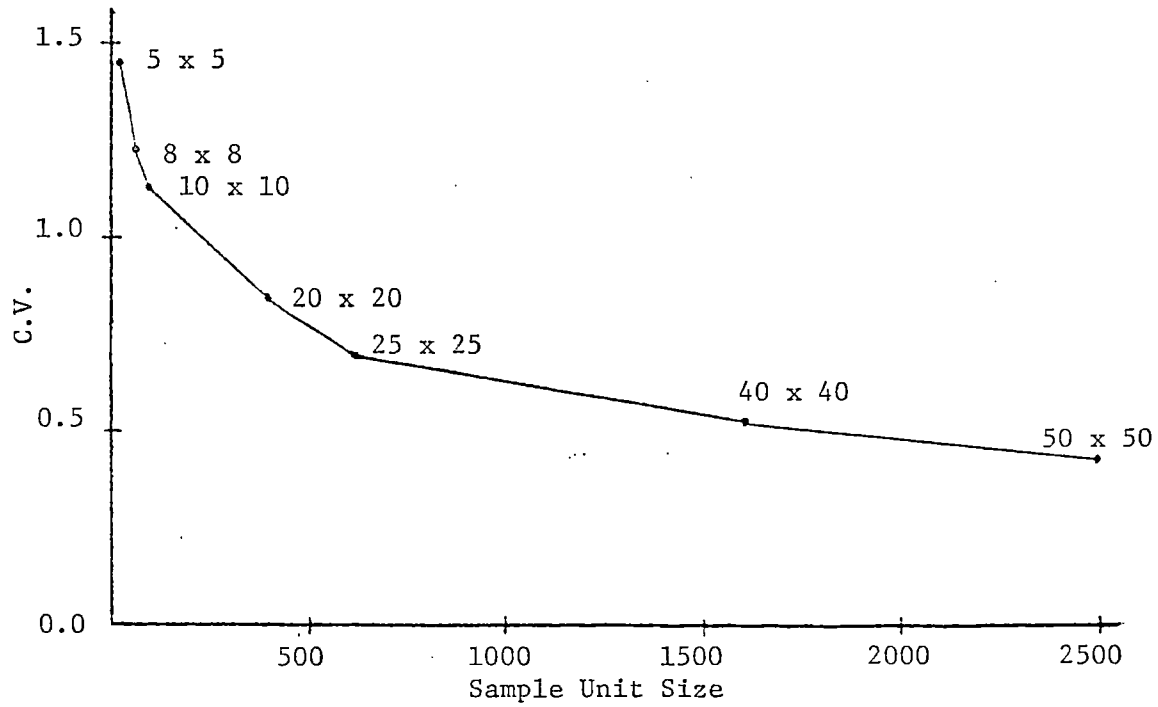


Figure 13. C.V. vs. sample unit size - MEADOW.

Although it was determined that square sample units would be used due to practical considerations (i.e., number of aerial photographs that need be photointerpreted to obtain ground truth measurements), the effect of different sample unit shape on coefficient of variation was investigated. This was done by holding the sample unit size constant (400 pixels for conifer) and determining the coefficient of variation for different shapes (Figure 14). It was seen that shapes other than square did in fact lower the coefficient of variation. This difference was not considered great enough to offset the practical advantages of square sample units, and thus square sample units were used throughout the experiment.

4.2 List Sampling Procedure

As mentioned in the previous chapter, determination of the sample size is based on the variance of the final estimate and the degree of accuracy desired. Since nothing is known about the variance at this point, preliminary samples were performed.

Preliminary samples were conducted to estimate the acreage of conifer and of meadow in the Middle Park study area. It was arbitrarily decided that five sample units would be used for the conifer sample and nine sample units for the meadow samples. List sampling was used to determine the specific sample units to use in the samples.

The first step in the conifer sample is to obtain a listing from the program PSU, showing the Landsat classification gridded off into sample units 20 pixels by 20 pixels in size (Table 7). This list shows sample unit number, location (lines and columns), x_i (number of pixels in sample unit classified as conifer) and the cumulative sum of x_i

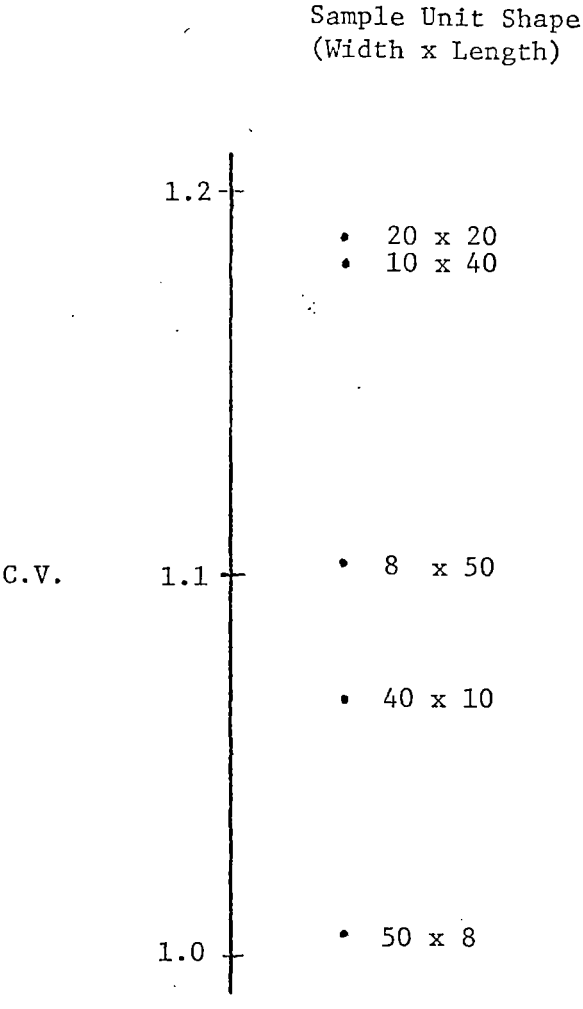


Figure 14. C.V. vs. sample unit shape - CONIFER.

Table 7. Sample list for CONIFER sample.

PRIMARY SAMPLE UNITS FOR CONIF PSU #		LINES			COLS.		X(I)	SUM X(I)
1	1	20	1	1	20	1	0.0	0.000
2	1	20	1	21	40	1	30.0	30.000
3	1	20	1	41	60	1	11.0	41.000
4	1	20	1	61	80	1	0.0	41.000
5	1	20	1	81	100	1	0.0	41.000
6	1	20	1	101	120	1	4.0	45.000
7	1	20	1	121	140	1	94.0	139.000
8	1	20	1	141	160	1	70.0	209.000
9	1	20	1	161	180	1	143.0	352.000
10	1	20	1	181	200	1	116.0	468.000
11	21	40	1	1	20	1	4.0	472.000
12	21	40	1	21	40	1	0.0	472.000
13	21	40	1	41	60	1	0.0	472.000
14	21	40	1	61	80	1	0.0	472.000
15	21	40	1	81	100	1	0.0	472.000
16	21	40	1	101	120	1	0.0	472.000
17	21	40	1	121	140	1	0.0	472.000
18	21	40	1	141	160	1	5.0	477.000
19	21	40	1	161	180	1	106.0	583.000
20	21	40	1	181	200	1	24.0	607.000
21	41	60	1	1	20	1	6.0	613.000
22	41	60	1	21	40	1	1.0	614.000
23	41	60	1	41	60	1	1.0	615.000
24	41	60	1	61	80	1	0.0	615.000
25	41	60	1	81	100	1	1.0	616.000
26	41	60	1	101	120	1	5.0	621.000
27	41	60	1	121	140	1	5.0	626.000
28	41	60	1	141	160	1	6.0	632.000
29	41	60	1	161	180	1	16.0	648.000
30	41	60	1	181	200	1	53.0	701.000
31	61	80	1	1	20	1	35.0	736.000
32	61	80	1	21	40	1	0.0	736.000
33	61	80	1	41	60	1	6.0	742.000
34	61	80	1	61	80	1	3.0	745.000
35	61	80	1	81	100	1	0.0	745.000
36	61	80	1	101	120	1	0.0	745.000
37	61	80	1	121	140	1	0.0	745.000
38	61	80	1	141	160	1	6.0	751.000
39	61	80	1	161	180	1	0.0	751.000
40	61	80	1	181	200	1	43.0	794.000
41	81	100	1	1	20	1	17.0	811.000
42	81	100	1	21	40	1	2.0	813.000
43	81	100	1	41	60	1	4.0	817.000
44	81	100	1	61	80	1	11.0	828.000
45	81	100	1	81	100	1	18.0	846.000
46	81	100	1	101	120	1	25.0	871.000
47	81	100	1	121	140	1	6.0	877.000
48	81	100	1	141	160	1	1.0	878.000
49	81	100	1	161	180	1	14.0	892.000
50	81	100	1	181	200	1	57.0	949.000
51	101	120	1	1	20	1	95.0	1044.000
52	101	120	1	21	40	1	12.0	1056.000
53	101	120	1	41	60	1	1.0	1057.000
54	101	120	1	61	80	1	0.0	1057.000
55	101	120	1	81	100	1	83.0	1140.000
56	101	120	1	101	120	1	147.0	1287.000
57	101	120	1	121	140	1	15.0	1302.000
58	101	120	1	141	160	1	5.0	1307.000
59	101	120	1	161	180	1	40.0	1347.000
60	101	120	1	181	200	1	153.0	1500.000
61	121	140	1	1	20	1	73.0	1573.000
62	121	140	1	21	40	1	93.0	1666.000
63	121	140	1	41	60	1	3.0	1669.000
64	121	140	1	61	80	1	18.0	1687.000
65	121	140	1	81	100	1	18.0	1705.000
66	121	140	1	101	120	1	127.0	1832.000
67	121	140	1	121	140	1	91.0	1923.000
68	121	140	1	141	160	1	15.0	1938.000
69	121	140	1	161	180	1	85.0	2023.000
70	121	140	1	181	200	1	124.0	2147.000

Table 7. (Continued)

PRIMARY SAMPLE UNITS FOR CONIF PSU =	LINES			COLS.			X(I)	SUM X(I)
71	141	160	1	1	20	1	191.0	2376.000
72	141	160	1	21	40	1	159.0	2537.000
73	141	160	1	41	60	1	64.0	2601.000
74	141	160	1	61	80	1	29.0	2630.000
75	141	160	1	81	100	1	2.0	2632.000
76	141	160	1	101	120	1	3.0	2635.000
77	141	160	1	121	140	1	190.0	2825.000
78	141	160	1	141	160	1	190.0	3015.000
79	141	160	1	161	180	1	143.0	3158.000
80	141	160	1	181	200	1	172.0	3330.000
81	161	180	1	1	20	1	121.0	3451.000
82	161	180	1	21	40	1	113.0	3564.000
83	161	180	1	41	60	1	190.0	3754.000
84	161	180	1	61	80	1	134.0	3888.000
85	161	180	1	81	100	1	27.0	3915.000
86	161	180	1	101	120	1	22.0	3937.000
87	161	180	1	121	140	1	36.0	3973.000
88	161	180	1	141	160	1	146.0	4119.000
89	161	180	1	161	180	1	236.0	4355.000
90	161	180	1	181	200	1	194.0	4549.000
91	181	200	1	1	20	1	89.0	4638.000
92	181	200	1	21	40	1	60.0	4698.000
93	181	200	1	41	60	1	177.0	4875.000
94	181	200	1	61	80	1	242.0	5117.000
95	181	200	1	81	100	1	213.0	5330.000
96	181	200	1	101	120	1	119.0	5449.000
97	181	200	1	121	140	1	110.0	5559.000
98	181	200	1	141	160	1	10.0	5569.000
99	181	200	1	161	180	1	40.0	5609.000
100	181	200	1	181	200	1	165.0	5774.000

MEAN = 57.74000
STANDARD DEVIATION = 66.42565
COEFFICIENT OF VARIATION = 1.15059

$(\sum_{j=1}^{i} x_j)$. Five random numbers between 0 and 5774 (the cumulative sum of the last sample unit) were then obtained using a Texas Instruments SR-51A hand calculator. The five random numbers selected were: 2061, 3459, 4039, 4319, and 4878. By using the list sampling procedure described in Chapter 3, the following sample units were thus selected for the sample: 70, 82, 88, 89, and 94.

For the meadow sample, a listing was obtained gridding off the study area into sample units 25 pixels by 25 pixels in size, and counting the number of pixels classified as meadow (Table 8). The nine random numbers between 0 and 4526 selected were: 308, 1110, 1156, 1485, 1983, 3098, 3396, 3866, and 4482. The corresponding sample units selected were: 4, 14, 15, 18, 25, 43, 49, 57, and 63.

In both of these samples only the ORIGINAL classification was used to perform the list sample. The sample units thus selected were used for all classifications.

For both conifer and meadow, then, five samples were made, corresponding to each of the five classifications used to determine selection probabilities. The selection probabilities will change between classifications, thus giving different acreage estimates for each classification. The purpose of this study was to determine if these differences are significant and, if so, if they are dependent on classification accuracy.

A second set of samples was performed for conifer. These samples took into account the fact that for a given random number sequence, because the values of the cumulative sum of x_i change, it is possible

Table 8. Sample list for MEADOW sample.

PRIMARY SAMPLE UNITS FOR PASTR								
POLE	LINES			COLS.			X(I)	SUM X(I)
1	1	25	1	1	25	1	54.0	54.000
2	1	25	1	26	50	1	58.0	112.000
3	1	25	1	51	75	1	42.0	154.000
4	1	25	1	76	100	1	275.0	429.000
5	1	25	1	101	125	1	28.0	457.000
6	1	25	1	126	150	1	96.0	553.000
7	1	25	1	151	175	1	56.0	609.000
8	1	25	1	176	200	1	69.0	678.000
9	26	50	1	1	25	1	74.0	752.000
10	26	50	1	26	50	1	85.0	837.000
11	26	50	1	51	75	1	3.0	840.000
12	26	50	1	76	100	1	154.0	994.000
13	26	50	1	101	125	1	77.0	1071.000
14	26	50	1	126	150	1	55.0	1126.000
15	26	50	1	151	175	1	33.0	1159.000
16	26	50	1	176	200	1	66.0	1225.000
17	51	75	1	1	25	1	105.0	1330.000
18	51	75	1	26	50	1	168.0	1498.000
19	51	75	1	51	75	1	56.0	1554.000
20	51	75	1	76	100	1	100.0	1654.000
21	51	75	1	101	125	1	144.0	1798.000
22	51	75	1	126	150	1	28.0	1826.000
23	51	75	1	151	175	1	37.0	1863.000
24	51	75	1	176	200	1	24.0	1887.000
25	76	100	1	1	25	1	130.0	2017.000
26	76	100	1	26	50	1	117.0	2134.000
27	76	100	1	51	75	1	41.0	2175.000
28	76	100	1	76	100	1	24.0	2199.000
29	76	100	1	101	125	1	42.0	2241.000
30	76	100	1	126	150	1	104.0	2345.000
31	76	100	1	151	175	1	89.0	2434.000
32	76	100	1	176	200	1	79.0	2513.000
33	101	125	1	1	25	1	48.0	2561.000
34	101	125	1	26	50	1	97.0	2658.000
35	101	125	1	51	75	1	32.0	2690.000
36	101	125	1	76	100	1	14.0	2704.000
37	101	125	1	101	125	1	16.0	2720.000
38	101	125	1	126	150	1	92.0	2812.000
39	101	125	1	151	175	1	56.0	2868.000
40	101	125	1	176	200	1	25.0	2893.000
41	126	150	1	1	25	1	61.0	2954.000
42	126	150	1	26	50	1	124.0	3078.000
43	126	150	1	51	75	1	66.0	3144.000
44	126	150	1	76	100	1	14.0	3158.000
45	126	150	1	101	125	1	47.0	3205.000
46	126	150	1	126	150	1	14.0	3219.000
47	126	150	1	151	175	1	76.0	3295.000
48	126	150	1	176	200	1	24.0	3319.000
49	151	175	1	1	25	1	78.0	3397.000
50	151	175	1	26	50	1	45.0	3442.000
51	151	175	1	51	75	1	97.0	3539.000
52	151	175	1	76	100	1	122.0	3661.000
53	151	175	1	101	125	1	86.0	3747.000
54	151	175	1	126	150	1	32.0	3779.000
55	151	175	1	151	175	1	18.0	3797.000
56	151	175	1	176	200	1	37.0	3834.000
57	176	200	1	1	25	1	173.0	4007.000
58	176	200	1	26	50	1	160.0	4167.000
59	176	200	1	51	75	1	22.0	4189.000
60	176	200	1	76	100	1	54.0	4243.000
61	176	200	1	101	125	1	66.0	4309.000
62	176	200	1	126	150	1	132.0	4441.000
63	176	200	1	151	175	1	59.0	4500.000
64	176	200	1	176	200	1	24.0	4524.000

LIBRARY COPY
ROCKY [illegible] RANGE
EXPERIMENT STATION

Table 8. (Continued)

MEAN = 70.71875
STANDARD DEVIATION = 49.27360
COEFFICIENT OF VARIATION = .69675

to select a different set of sample units from different classifications in using the same list of random numbers.

To investigate the effect of this, list samples were conducted for the four additional classifications using the same random number list used to list sample the original classification. The interval for possible random numbers is different for different classifications ((0, 5774] for the original classification, (0,8708] for JOSESIGS, etc.) due to the fact that the cumulative sum of x_i is different. Therefore, the original random number list was normalized to values between 0 and 1 by dividing by the cumulative sum of x_i for the last sample unit of the original classification (5774). For the four additional classifications, then, these normalized numbers were multiplied by the value of the cumulative sum of x_i for the last sample unit of the corresponding classification (8708 for JOSESIGS, 5509 for SCRAMBL1, 5039 for SCRAMBL2, and 6637 for SCRAMBL3). These calculations are shown in Table 9. The resulting random number lists were then used to list sample from the additional classifications. The listings from the program PSU for each classification used to perform the list samples are given in Tables 25 through 34, Appendix B.

For the classification SCRAMBL1 the sample units selected were the same as the original sample. For JOSESIGS, SCRAMBL2, and SCRAMBL3, however, the set of sample units selected for the sample was different. Thus, in addition to the effect the different selection probabilities has on acreage estimate calculations, the effect of the possibility of selecting different sample units from different classifications was studied.

Table 9. Calculations for new random number lists.

r = original random number

r' = normalized random number = r/5774

$r'' = r' \cdot \sum_{i}^{N} x_i$ (new classification)

SU = sample unit selected by list sampling

Classif.	$\sum_{i}^{N} x_i$	r	r'	r''	SU
ORIGINAL	5774	2061	.357	2061	70
		3459	.599	3459	82
		4039	.700	4039	88
		4319	.748	4319	89
		4878	.845	4878	94
JOSESIGS	8708		.357	3108.8	67
			.599	5216.1	80
			.700	6095.6	84
			.748	6513.6	89
			.845	7358.3	93
SCRAMBL1	5509		.357	1966.6	70
			.599	3299.9	82
			.700	3856.3	88
			.748	4120.7	89
			.845	4655.1	94
SCRAMBL2	5039		.357	1798.9	69
			.599	3018.4	81
			.700	3527.3	88
			.748	3769.2	89
			.845	4258.0	93
SCRAMBL3	6637		.357	2369.4	67
			.599	3975.6	80
			.700	4645.9	86
			.748	4964.5	89
			.845	5608.3	93

4.3 Obtaining Sample Measurements from Aerial Photography

Once the specific sample units to be included in the samples have been determined, the next step is to obtain the ground truth measurements to be used in the acreage estimate calculations. The ground truth information was obtained primarily from 1:24,000 scale color infrared aerial photography. The photography was taken during a United States Forest Service overflight of June 28 and 29, 1977 (U.S.F.S. Experiment 4151-2 Mission FS-3). In addition, CIR aerial photography at scales of 1:65,000 and 1:130,000 (U.S.F.S. U-2 Mission 76-096 taken June 28, 1976) was used where necessary to fill in the flight line gaps of the 1:24,000 photography.

Photointerpretation was performed on the aerial photography for the entire study area (less flight line gaps where there was no coverage). Since the photography consisted of overlapping stereo pairs, only every other photograph in a flight line was interpreted. The interpretation was performed by delineating the two land cover classes of interest, conifer and meadow, on an acetate overlay attached to the photographic transparencies.

This information was then used to prepare a ground truth map for the study area. The base maps used were the four topographic maps covering the Middle Park study area (Kremmling, Hot Sulphur Springs, Mt. Powell, and Ute Peak quadrangles) at a scale of 1:62,500. This was done by using a Bausch and Lomb Zoom Transfer Scope to transfer the information from the aerial photographs to the topographic base maps.

Inaccuracies arose in the preparation of the ground truth maps, due primarily to difficulty in aligning the aerial photographs with the topographic base maps. Alignment of the two images is done by

locating features which are easily identified on both the photography and the map. However, two of the topographic maps are so old (prepared in 1933) that many of the identifiable features, such as roads and reservoirs, that are found on the map are not found on the photography and vice versa. In addition, the field of view of the Zoom Transfer Scope in this situation was so small that in many of the undeveloped areas no roads, water bodies, etc. could be found with which to allign the two images. These inaccuracies no doubt made the final acreage estimates less accurate than they otherwise would have been. However, the effect on the results of this experiment were considered to be minimal, since the effect would be about the same for each classification.

An attempt was made to obtain ground truth data for the entire study area, but this was found to be impossible due to gaps between the flight lines in the aerial photography. Additional aerial photography (the 1:65,000 and 1:130,000 photography) was used in an attempt to fill in these gaps, but very little could be added to the original ground truth data due to gaps and severe cloud coverage in the photography. At later stages in the experiment, where additional sample units were randomly selected, some of these sample units could not be used due to the incompleteness of the ground truth data and new sample units had to be selected. The unusable sample units are distributed randomly rather than systematically with respect to the vegetation patterns, and thus the integrity of the random selection procedure was not considered to be severely effected.

Ground truth measurements for the sample units for the conifer samples and the meadow sample are given in Tables 10 and 11.

Table 10. Ground truth area measurements.

CONIFER

Sample Unit	Lines	Columns	Area ($mi.^2$)
19	21-40	161-180	1.431
40	61-180	181-200	0.615
51	101-120	1-20	3.236
59	101-120	161-180	0.599
60	101-120	181-200	2.714
62	121-140	21-40	2.573
67	121-140	121-140	3.152
69	121-140	161-180	1.694
70	121-140	181-200	2.877
80	141-160	181-200	3.189
81	161-180	1-20	2.302
82	161-180	21-40	2.425
83	161-180	41-60	3.087
84	161-180	61-80	1.912
86	161-180	101-120	0.127
88	161-180	141-160	2.123
89	161-180	161-180	3.845
93	181-200	41-60	3.394
94	181-200	61-80	4.135

Table 11. Ground truth area measurements.

MEADOW

Sample Unit	Lines	Columns	Area ($mi.^2$)
4	1-25	76-100	2.948
14	26-50	126-150	0.618
15	26-50	151-175	0.084
18	51-75	26-50	2.168
25	76-100	1-25	0.340
43	126-150	51-75	0.900
49	151-175	1-25	0.357
57	176-200	1-25	1.117
63	176-200	151-175	0.704

4.4 Calculations and Results

The experiment described above can be divided into three separate experiments:

1. A preliminary sample for conifer acreage estimation using the same five sample units for each classification
2. A preliminary sample for conifer acreage estimation using the same list of random numbers to select five sample units which are possibly different for each classification.
3. A preliminary sample for meadow acreage estimation using the same nine sample units for each classification.

Based on the results of these experiments, two additional experiments were performed:

4. Four additional samples for conifer acreage estimation using five sample units in each sample and using the same sample units for each classification.
5. A final sample for conifer acreage estimation using all nineteen sample units from previous samples and using the same sample units for each classification.

An estimate of the acreage of conifer or meadow is given by equation 3.6:

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i} \tag{3.6}$$

Sample variance and standard deviation are given by equations 3.7 through 3.9:

$$\text{sample variance} = s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (A_i - \hat{A})^2 \tag{3.7}$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} A_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} A_i\right)^2\right] \tag{3.8}$$

where $A_i = \frac{y_i}{p_i}$

$$\text{sample standard deviation} = s = \sqrt{s^2} \tag{3.9}$$

As stated in Chapter 3 the true variance of the estimate, $\hat{A}$, is:

$$\text{Var}(\hat{A}) = \frac{1}{n}\sum_{i=1}^{n} p_i\,(A_i - A)^2 \tag{3.10}$$

An unbiased estimator of this variance is:

$$\hat{\text{Var}}(\hat{A}) = \frac{1}{n(n-1)}\sum_{i=1}^{n} (A_i - \hat{A})^2 = s^2/n \tag{3.11}\ (3.12)$$

Estimators of the population standard deviation and coefficient of variation are:

$$\hat{\sigma} = \sqrt{\hat{\text{Var}}(\hat{A})} = s/\sqrt{n} \tag{3.13}$$

$$\text{C.V.} = \hat{\sigma}/\hat{A} \tag{3.14}$$

Based on this coefficient of variation, the sample size needed to achieve a specified degree of accuracy is computed by equations 3.15 and 3.16:

$$n' = \frac{t^2\ \text{C.V.}^2}{\text{A.E.}^2} \tag{3.15}$$

$$n = \frac{n'}{1 + \frac{n'}{N}} \tag{3.16}$$

For the three preliminary sample experiments, five acreage estimates are obtained corresponding to each of the five classifications. By computing the coefficients of variation for these estimates it is determined whether further sampling need be conducted to achieve the level of accuracy specified.

For each classification the following information is recorded: sample size, n; population size, N; and cumulative sum over the entire population, $\sum_{i=1}^{N} x_i$. For each sample unit in a sample, the following are recoreded: number of pixels in the sample unit of the class of interest, x_i; selection probability, $p_i = x_i / \sum_{i=1}^{N} x_i$; ground truth acreage measurement, y_i; and the total acreage estimate provided by that sample unit, $A_i = y_i/p_i$. The data recoreded and calculations made are summarized for each experiment as follows.

Experiment 1

In this experiment five sets of sample calculations, corresponding to each of the five classifications, were performed to estimate the acreage of conifer in the Middle Park study area. The data for these classifications is shown in Table 12.

The calculations for the classification ORIGINAL, performed on a Texas Instruments SR51-II hand calculator, are given below as an example.

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i} = \frac{1}{5} \sum_{i=1}^{5} A_i$$

$$= \frac{1}{5} (95.5 + 123.9 + 84.0 + 94.1 + 98.7)$$

$$= 99.2$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (A_i - \hat{A})^2 = \frac{1}{n-1} \left[\sum_{i=1}^{n} A_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} A_i \right)^2 \right]$$

$$= \frac{1}{4} [95.5^2 + 123.9^2 + 84.0^2 + 94.1^2 + 98.7^2 - \frac{1}{5} (496.2)^2]$$

$$= 220.8$$

Table 12. Experiment 1 sample data.

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	174	.03014	2.877	95.5
82	113	.01957	2.425	123.9
88	146	.02529	2.123	84.0
89	236	.04087	3.845	94.1
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 99.2 $s^2 = 220.8$ $s = 14.9$

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	225	.02584	2.877	111.3
82	205	.02354	2.425	103.0
88	182	.02090	2.123	101.6
89	288	.03307	3.845	116.3
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 113.1 $s^2 = 164.5$ $s = 12.8$

Table 12. (Continued)

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	165	.02995	2.877	96.1
82	111	.02015	2.425	120.4
88	140	.02541	2.123	83.5
89	223	.04048	3.845	95.0
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 98.6 $s^2 = 179.9$ $s = 13.4$

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	148	.02941	2.877	97.8
82	101	.02007	2.425	120.8
88	120	.02385	2.123	89.0
89	188	.03736	3.845	102.9
94	202	.04014	4.135	103.0

$\hat{A}$ (mi.^{2}) = 102.7 $s^2 = 134.8$ $s = 11.6$

Table 12. (Continued)

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	P_i	y_i	A_i (mi.^{2})
70	175	.02637	2.877	109.1
82	131	.01974	2.425	122.9
88	151	.02275	2.123	93.3
89	242	.03646	3.845	105.5
94	240	.03616	4.135	114.4

$\hat{A}$ (mi.^{2}) = 109.0 $s^2 = 119.8$ $s = 10.9$

$$s = \sqrt{s^2} = \sqrt{220.8} = 14.9$$

$$\hat{Var}(\hat{A}) = s^2/n = 220.8/5 = 44.2$$

$$\hat{\sigma} = \sqrt{\hat{Var}(\hat{A})} = \sqrt{44.2} = 6.6$$

$$C.V. = \hat{\sigma}/\hat{A} = 6.6/99.2 = .067 \text{ or } 6.7\%$$

If an allowable error of 10% is specified and the t statistic for n-1 = 4 degrees of freedom and 90% confidence level is used, the necessary sample size is:

$$n' = \frac{t^2\ C.V.^2}{A.E.^2} = \frac{(2.132)^2\ (.067)^2}{(.10)^2} = 2.04$$

$$n = \frac{n'}{1 + \frac{n'}{N}} = \frac{2.04}{1 + \frac{2.04}{100}} = 2.00 \simeq 2$$

Thus a sample consisting of two sample units is needed for the specified level of accuracy. Since the preliminary sample consists of five sample units, no further sampling need be done, and this can be called the final sample.

The same set of calculations was performed for the other four classifications. These are summarized in Table 13 for all five classifications.

A good graphic comparison of the results for each classification is given by a plot of the estimated acreage and 90% confidence interval of the estimate. The calculation of the 90% confidence interval is:

$$\hat{A} \pm ts/\sqrt{n}$$

Table 13. Sample summary.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

$\hat{V}ar(\hat{A})$ = estimate of population variance = s^2/n

$\hat{\sigma}$ = estimate of population standard deviation = $\sqrt{\hat{V}ar(\hat{A})}$

C.V. = coefficient of variation = $\hat{\sigma}/\hat{A}$

Experiment: 1 Land Cover Class: CONIFER

Classification	$\hat{A}$	s	$\hat{V}ar(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	99.2	14.9	44.4	6.7	.067
JOSESIGS	113.1	12.8	32.8	5.7	.051
SCRAMBL1	98.6	13.4	35.9	6.0	.061
SCRAMBL2	102.7	11.6	26.9	5.2	.051
SCRAMBL3	109.0	10.9	23.8	4.9	.045

where: $\hat{A}$ = acreage estimate

t = tabular t value at 90% confidence level and n-1 degrees of freedom

s = sample standard deviation

n = sample size

This plot is shown in Figure 15. From this it can be seen that the estimates are all fairly close, at least to the point that the confidence intervals overlap. A detailed discussion of these results is given in the next chapter.

Experiment 2

This experiment is similar to the previous experiment to estimate the acreage of conifer in the study area. The difference is that rather than simply using the same sample units with each classification, a list sample was conducted for each classification using the same set of random numbers (normalized) to select sample units. Because the values of the cumulative sum of x_i are different for each classification, this raises the possibility of selecting different sample units for use with each classification.

The samples were the same as in the previous experiment for the classification ORIGINAL (by design) and SCRAMBL1 (by chance). The JOSESIGS and SCRAMBL3 samples had one sample unit in common with the samples of the previous experiment, while the SCRAMBL2 sample had two sample units in common.

The data for this experiment is given in Table 36, Appendix C. A summary of the calculations is shown in Table 14. A plot of the means

90% confidence interval = $\hat{A} \pm ts/\sqrt{n}$

t (4 d.f., 90% confidence level) = 2.132 n = 5

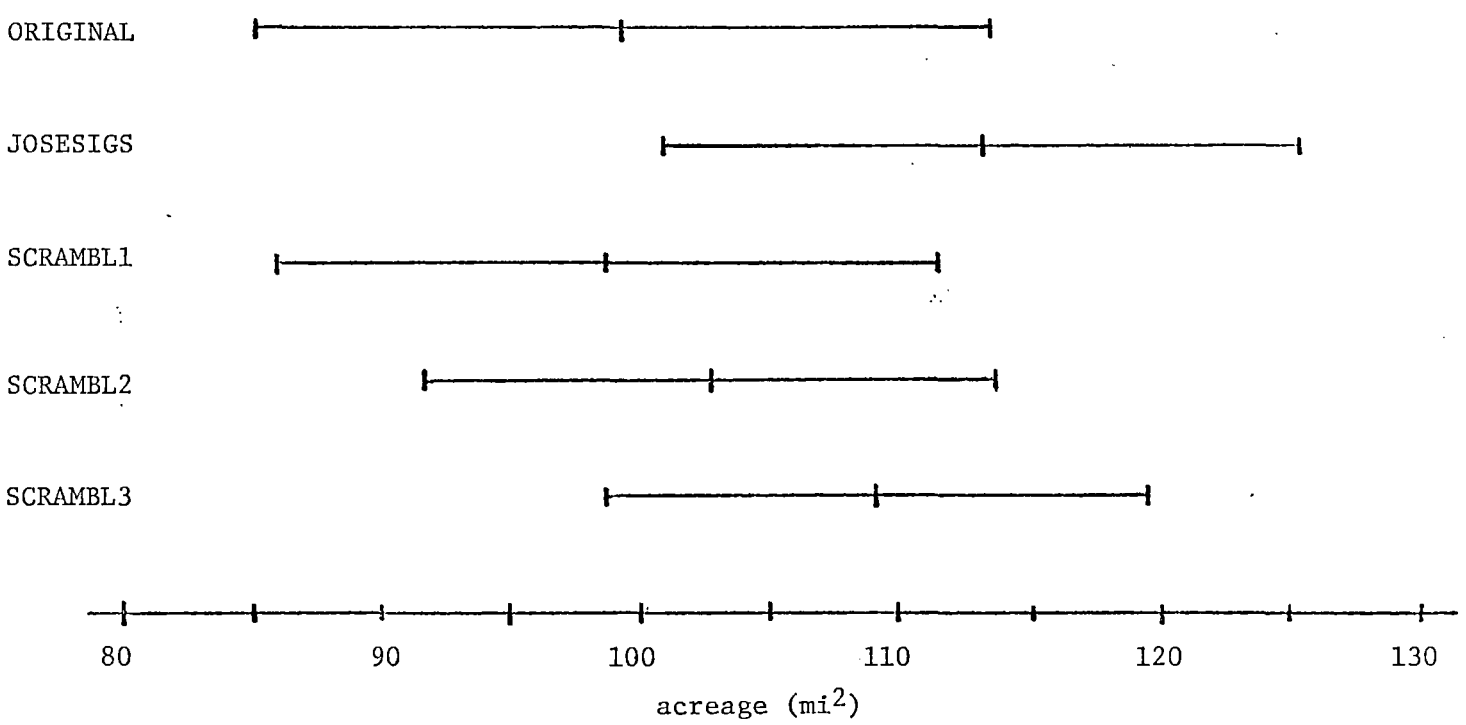


Figure 15. 90% confidence intervals for acreage estimates - Experiment 1.

Table 14. Sample summary.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

$\hat{Var}(\hat{A})$ = estimate of population variance = s^2/n

$\hat{\sigma}$ = estimate of population standard deviation = $\sqrt{\hat{Var}(\hat{A})}$

C.V. = coefficient of variation = $\hat{\sigma}/A$

Experiment: 2 Land Cover Class: CONIFER

Classification	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	99.2	14.9	44.4	6.7	.067
JOSESIGS	113.4	14.6	42.6	6.5	.058
SCRAMBL1	98.6	13.4	35.9	6.0	.061
SCRAMBL2	108.6	12.9	33.3	5.8	.053
SCRAMBL3	109.4	55.3	611.6	24.7	.226

and 90% confidence intervals for each classification is shown in Figure 16.

One easily noticed result of this experiment is that the variance (and thus the 90% confidence interval) of the sample using SCRAMBL3 is quite a bit larger than for any of the other samples of this or the previous experiment. This effect is discussed in detail in the next chapter.

Experiment 3

This experiment consisted of a preliminary sample using each of five classifications to estimate the acreage of meadow in the study area. Nine sample units were used in the sample (rather than five units as in the conifer preliminary sample), because ground truth data was known for nine randomly selected sample units before the experiment was conducted.

The data for this experiment is given in Table 37, Appendix C. The calculations were the same as for the previous experiments. The results are summarized in Table 15. A plot of means and 90% confidence intervals is given in Figure 17.

The results show that the relative variances of these samples (as shown by the coefficients of variation) are several magnitudes larger than for the conifer samples. This is reflected in the calculation of the sample size needed for an allowable error of 10%. Using C.V. = .293 from the classification ORIGINAL and t = 2.132:

$$n' = \frac{t^2 \, C.V.^2}{A.E.^2} = \frac{(2.132)^2 \, (.293)^2}{(.10)^2} = 39.02$$

90% confidence interval = $\hat{A} \pm ts/\sqrt{n}$

t (4 d.f., 90% confidence level) = 2.132 n = 5

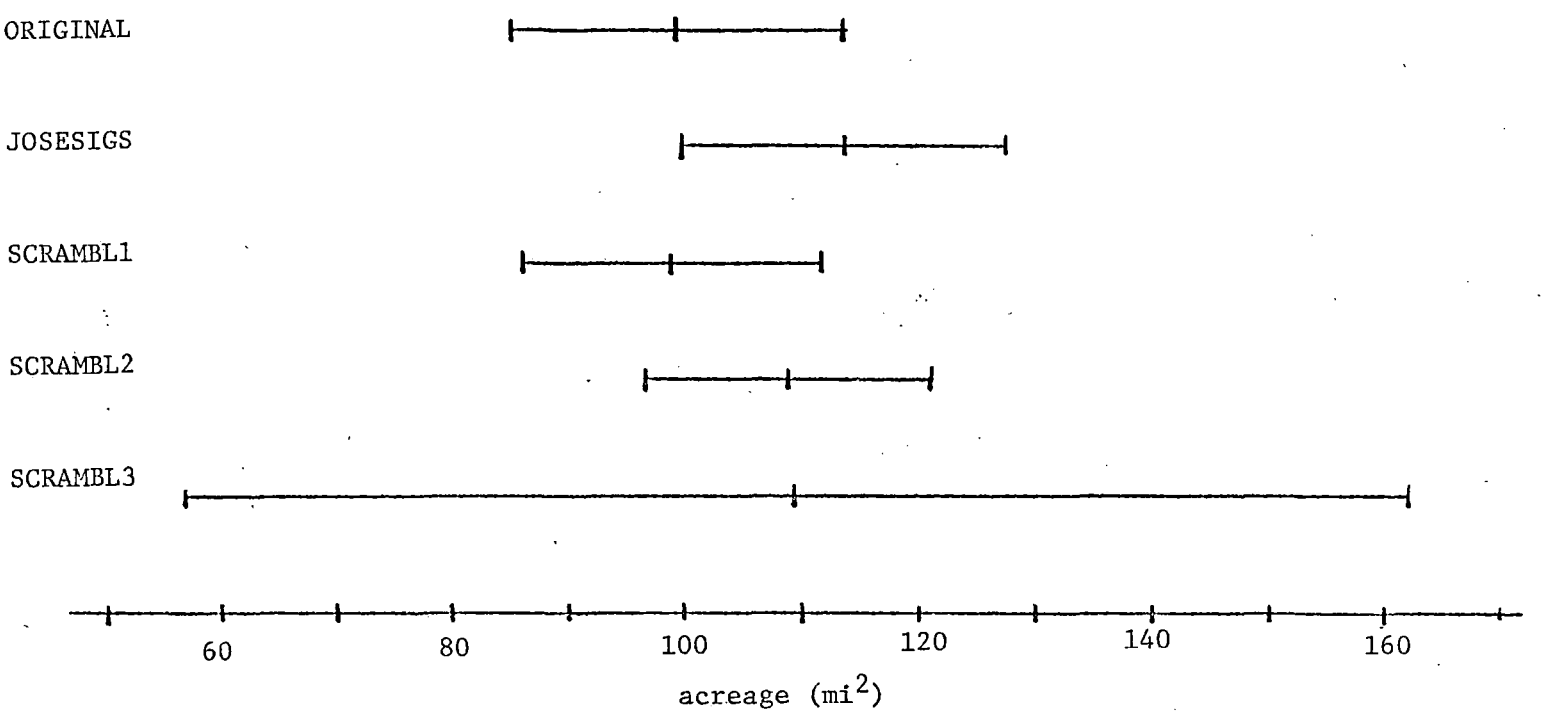


Figure 16. 90% confidence intervals for acreage estimates - Experiment 2.

Table 15. Sample summary.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

$\hat{Var}(\hat{A})$ = estimate of population variance = s^2/n

$\hat{\sigma}$ = estimate of population standard deviation = $\sqrt{\hat{Var}(\hat{A})}$

C.V. = coefficient of variation = $\hat{\sigma}/\hat{A}$

Experiment: 3 Land Cover Class: MEADOW

Classi-fication	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	38.5	20.2	45.3	6.7	.175
JOSESIGS	39.0	23.0	58.8	7.7	.197
SCRAMBL1	38.7	20.4	46.2	6.8	.176
SCRAMBL2	42.7	24.7	67.8	8.2	.193
SCRAMBL3	41.1	22.1	54.3	7.4	.179

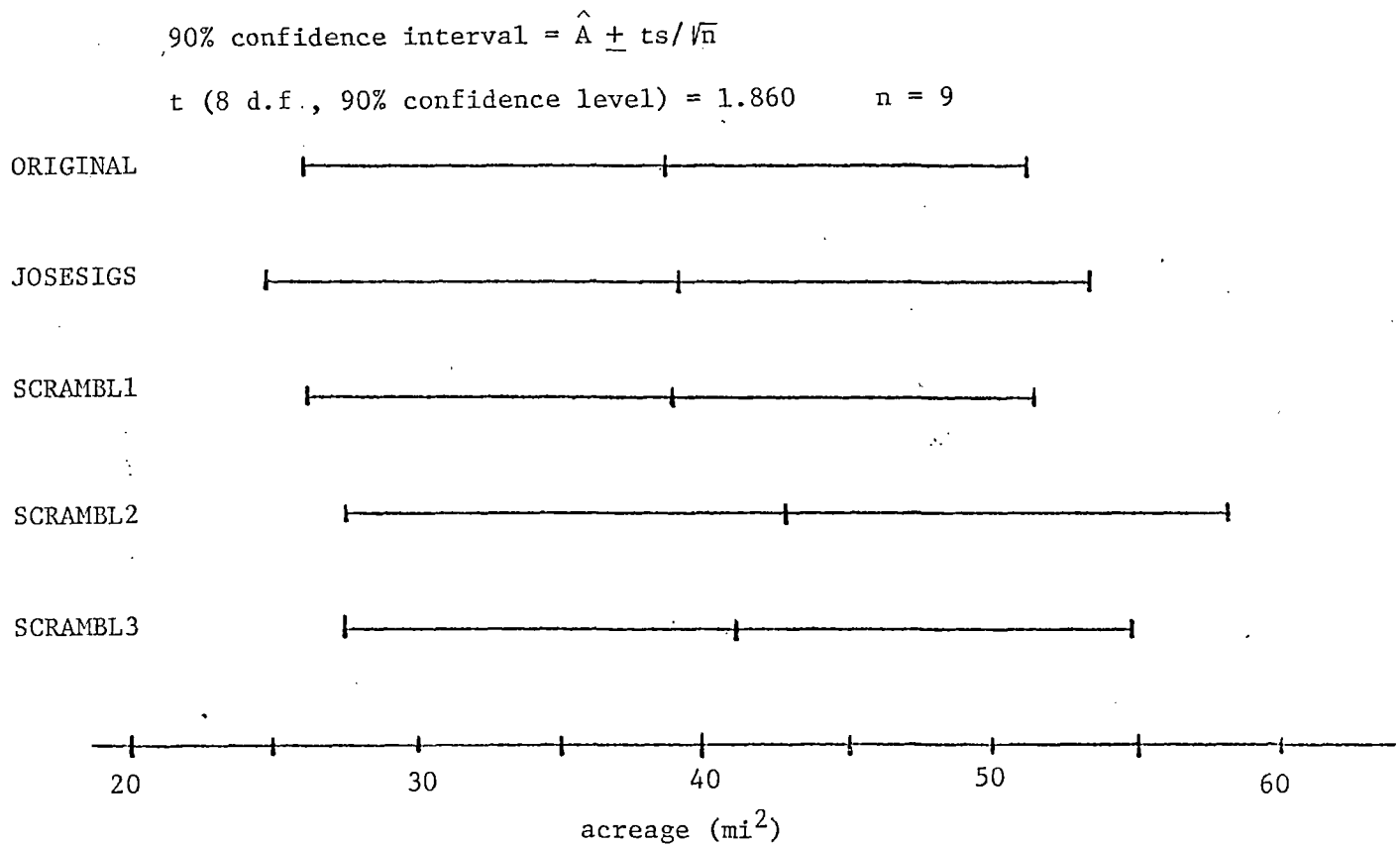


Figure 17. 90% confidence intervals for acreage estimates - Experiment 3.

$$n = \frac{n'}{1 + \frac{n'}{N}} = \frac{39.02}{1 + \frac{39.02}{64}} = 24.24$$

This value is rounded up, so that the necessary sample size is 25 sample units. This would mean sampling 40% of the population. The conclusion was drawn from these results that sampling just does not work very well for this particular land cover class. Therefore, all further sampling and analysis was done for the conifer class only. An explanation for the results observed for meadow is proposed in the next chapter.

Experiment 4

In this experiment five separate samples were conducted with each classification to estimate the conifer acreage. One sample was that used in Experiment 1. The other four samples were selected by obtaining new lists of five random numbers. For a given sample, the same sample units were used with each classification. As was to be expected, there was some duplication of sample units between samples, particularly for those sample units with a high selection probability.

The purpose of this experiment was to provide information beyond that supplied by just one sample. The increased amount of data could help in isolating random effects from the systematic effects due to the particular classification used, and provide data for use with tests of the significance of differences between classification estimates.

The data for the five samples is given in Table 38, Appendix C. The results of the five samples are summarized in Table 16. A plot of the means and 90% confidence intervals are given in Figure 18. A summary of the results showing acreage estimates and sample standard

Table 16. Sample summary.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

$\hat{Var}(\hat{A})$ = estimate of population variance = s^2/n

$\hat{\sigma}$ = estimate of population standard deviation = $\sqrt{\hat{Var}(\hat{A})}$

C.V. = coefficient of variation = $\hat{\sigma}/A$

Experiment: 4-1 Land Cover Class: CONIFER

Classification	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	99.2	14.9	44.4	6.7	.067
JOSESIGS	113.1	12.8	32.8	5.7	.051
SCRAMBL1	98.6	13.4	35.9	6.0	.061
SCRAMBL2	102.7	11.6	26.9	5.2	.051
SCRAMBL3	109.0	10.9	23.8	4.9	.045

Table 16. (Continued)

Experiment: 4-2 Land Cover Class: CONIFER

Classi-fication	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	90.5	8.3	13.8	3.7	.041
JOSESIGS	114.2	11.0	24.2	4.9	.043
SCRAMBL1	90.8	8.7	15.1	3.9	.043
SCRAMBL2	93.0	10.5	22.1	4.7	.050
SCRAMBL3	99.9	12.4	30.8	5.5	.056

Experiment: 4-3 Land Cover Class: CONIFER

Classi-fication	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	137.7	56.9	647.5	25.4	.185
JOSESIGS	127.7	14.3	40.9	6.4	.050
SCRAMBL1	136.9	56.0	627.2	25.0	.183
SCRAMBL2	132.5	47.1	443.7	21.1	.159
SCRAMBL3	135.8	41.3	341.1	18.5	.136

Table 16. (Continued)

Experiment: 4-4 Land Cover Class: CONIFER

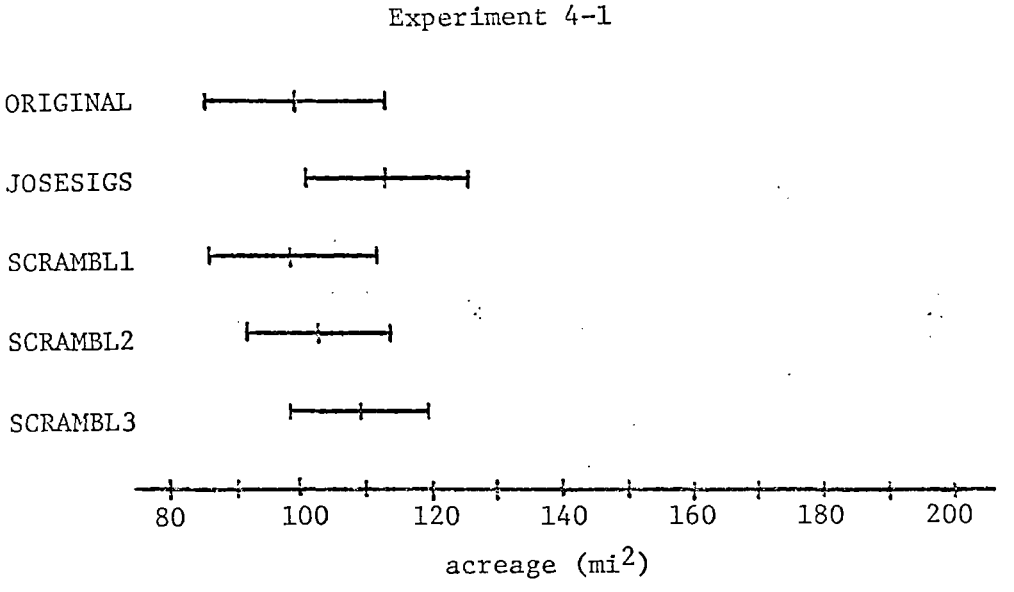
Classi-fication	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	98.5	9.7	18.8	4.3	.044
JOSESIGS	110.9	1.6	0.5	0.7	.006
SCRAMBL1	99.0	9.4	17.7	4.2	.042
SCRAMBL2	103.0	15.1	45.6	6.8	.066
SCRAMBL3	103.6	11.8	27.8	5.3	.051

Experiment: 4-5 Land Cover Class: CONIFER

Classi-fication	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	113.4	28.5	162.5	12.7	.112
JOSESIGS	115.4	18.1	65.5	8.1	.070
SCRAMBL1	111.9	27.9	155.7	12.5	.112
SCRAMBL2	113.2	25.7	132.1	11.5	.102
SCRAMBL3	116.6	26.1	136.2	11.7	.100

90% confidence interval = $\hat{A} \pm ts/\sqrt{n}$

t (4 d.f., 90% confidence level) = 2.132 n = 4



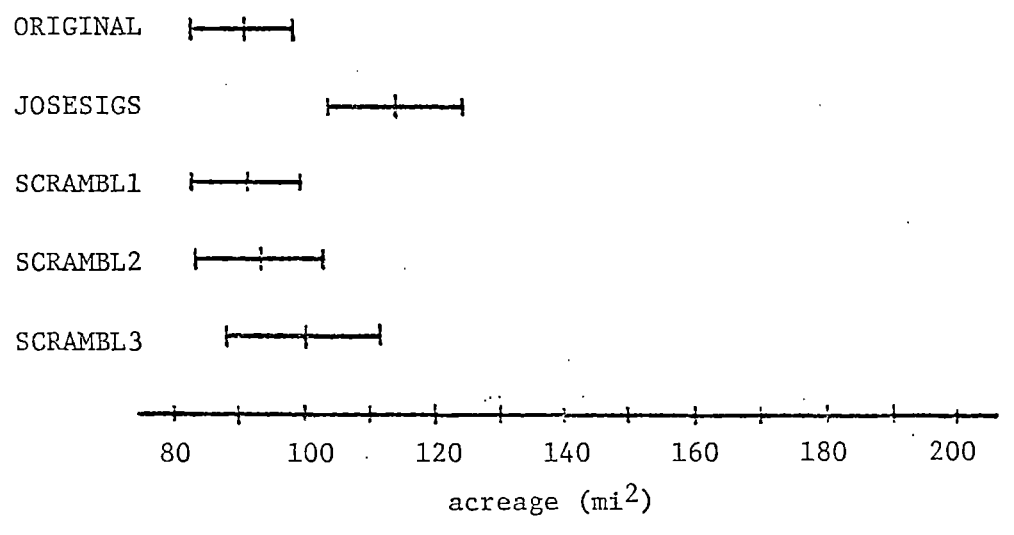


Figure 18. 90% confidence intervals for acreage estimates - Experiment 4.

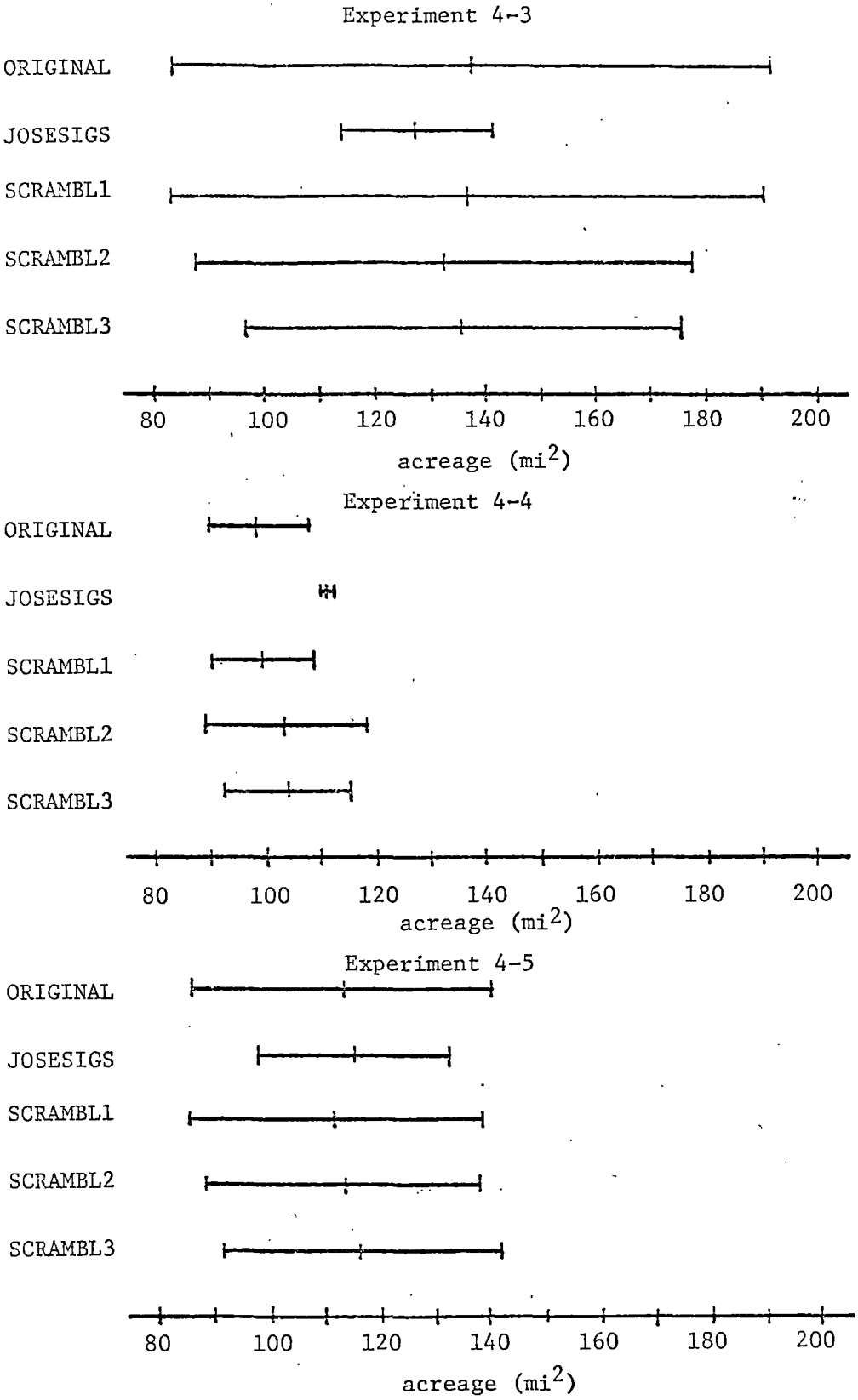


Figure 18. (Continued)

deviation by classification and by sample is given in Table 17. A plot of the means and 90% confidence intervals for all samples over all classifications is given in Figure 19.

As can be seen from Figure 19, there is quite a bit of variability between the means and variances of the different samples, even within one classification. Several tests of comparison of means and variances were performed. These tests were not strictly valid, however, because the samples were not independent of each other, i.e., different samples had sample units in common (Dixon, 1978). In any case, these tests merely showed what can be inferred by viewing the plots of means and confidence intervals--that there is quite a bit more variability between samples than between classifications.

The conclusion drawn from this result was that more sample units were required in the samples. The calculation for the sample size necessary for a 10% allowable error from the first sample showed that only two sample units were needed. However, succeeding samples (samples 3 and 5 of this experiment) showed a much greater sample variance. As a worst case test, sample 3 using the ORIGINAL classification shows the following coefficient of variation:

$$\hat{\sigma}/\hat{A} = 25.4/137.7 = .184$$

The calculation of necessary sample size is:

$$n' = \frac{t^2 \ C.V.^2}{A.E.^2} = \frac{(2.132)^2 \ (.184)^2}{(.10)^2} = 15.39$$

$$n = \frac{n'}{1 + \frac{n'}{N}} = \frac{15.39}{1 + \frac{15.39}{100}} = 13.34 \simeq 14 \text{ (rounding up)}$$

Conifer

For original

$n = 19$ 100

C.V. = .085

$t_{n-1} =$ 1.734

$$n' = \frac{t^2 CV^2}{(.10)^2} = \frac{(1.734)^2(.085)^2}{(.10)^2} = 2.17$$

$$\frac{1 + 2.17}{1 + \frac{2.17}{100}} \simeq 3$$

95% conf limit, 5% error

$t = 2.101$

$$\frac{(2.101)^2 (.085)^2}{(.05)^2} = 12.76$$

$$\frac{12.76}{1 + \frac{12.76}{100}} = 11.32 \rightarrow 12$$

Table 17. Experiment 4 - summary of all samples.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

Sample		Classif. ORIGINAL	JOSESIGS	SCRAMBL1	SCRAMBL2	SCRAMBL3
1	$\hat{A}$	99.2	113.1	98.6	102.7	109.0
	s	14.9	12.8	13.4	11.6	10.9
2	$\hat{A}$	90.5	114.2	90.8	93.0	99.9
	s	8.3	11.0	8.7	10.5	12.4
3	$\hat{A}$	137.7	127.7	136.9	132.5	135.8
	s	56.9	14.3	56.0	47.1	41.3
4	$\hat{A}$	98.5	110.9	99.0	103.0	103.6
	s	9.7	1.6	9.4	15.1	11.8
5	$\hat{A}$	113.4	115.4	111.9	113.2	116.6
	s	28.5	18.1	27.9	25.7	26.1

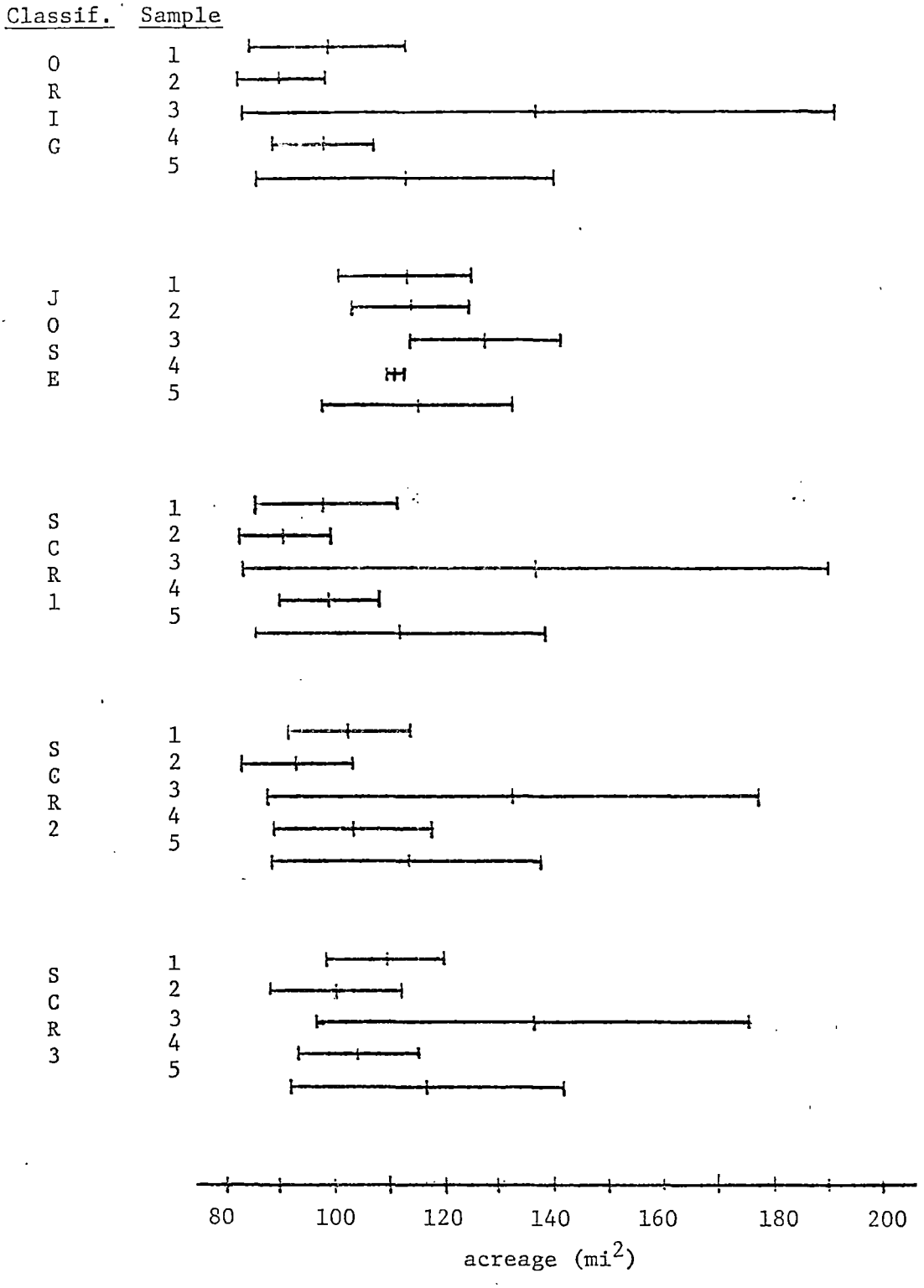


Figure 19. 90% confidence intervals for acreage estimates - Experiment 4.

Since a total of 16 different sample units were used in the 5 samples of this experiment, a single sample using all 16 sample units, plus 3 additional sample units from Experiment 2, was performed using each classification. This was done in Experiment 5, described below.

Experiment 5

This experiment estimated acreage of conifer in the study area by performing a single sample of 19 sample units using each classification. The same 19 sample units were used with each classification and were those sample units from previous experiments for which ground truth measurements had been obtained. This experiment is the same as Experiment 1 except for the sample size.

The data for these samples are given in Table 39, Appendix C. A summary of the results is shown in Table 18 and a plot of the means and 90% confidence intervals in Figure 20.

A test of the equality of acreage estimates obtained using the five classifications was performed. This involved using a t test for equality of means (Guenther, 1965). Ten tests were performed to assess the equality of all combinations of acreage estimates. The ten hypotheses tested (where subscript denotes classification) were:

1) H_0: $\mu_1 = \mu_2$, vs. H_1: $\mu_1 \neq \mu_2$
2) HO: $\mu_1 = \mu_3$, vs. H_1: $\mu_1 \neq \mu_3$
3) H_0: $\mu_1 = \mu_4$, vs. H_1: $\mu_1 \neq \mu_4$
4) H_0: $\mu_1 = \mu_5$, vs. H_1: $\mu_1 \neq \mu_5$
5) H_0: $\mu_2 = \mu_3$, vs. $_1$: $\mu_2 \neq \mu_3$

Table 18. Sample summary.

$\hat{A}$ = sample acreage estimate (mi^2)

s = sample standard deviation

$\hat{Var}(\hat{A})$ = estimate of population variance = s^2/n

$\hat{\sigma}$ = estimate of population standard deviation = $\sqrt{\hat{Var}(\hat{A})}$

C.V. = coefficient of variation = $\hat{\sigma}/A$

Experiment: 5 Land Cover Class: CONIFER

Classification	$\hat{A}$	s	$\hat{Var}(\hat{A})$	$\hat{\sigma}$	C.V.
ORIGINAL	108.1	40.0	84.2	9.2	.085
JOSESIGS	110.1	25.1	33.2	5.8	.052
SCRAMBL1	107.6	39.8	83.4	9.1	.085
SCRAMBL2	108.2	36.3	69.4	8.3	.077
SCRAMBL3	111.4	34.9	64.1	8.0	.072

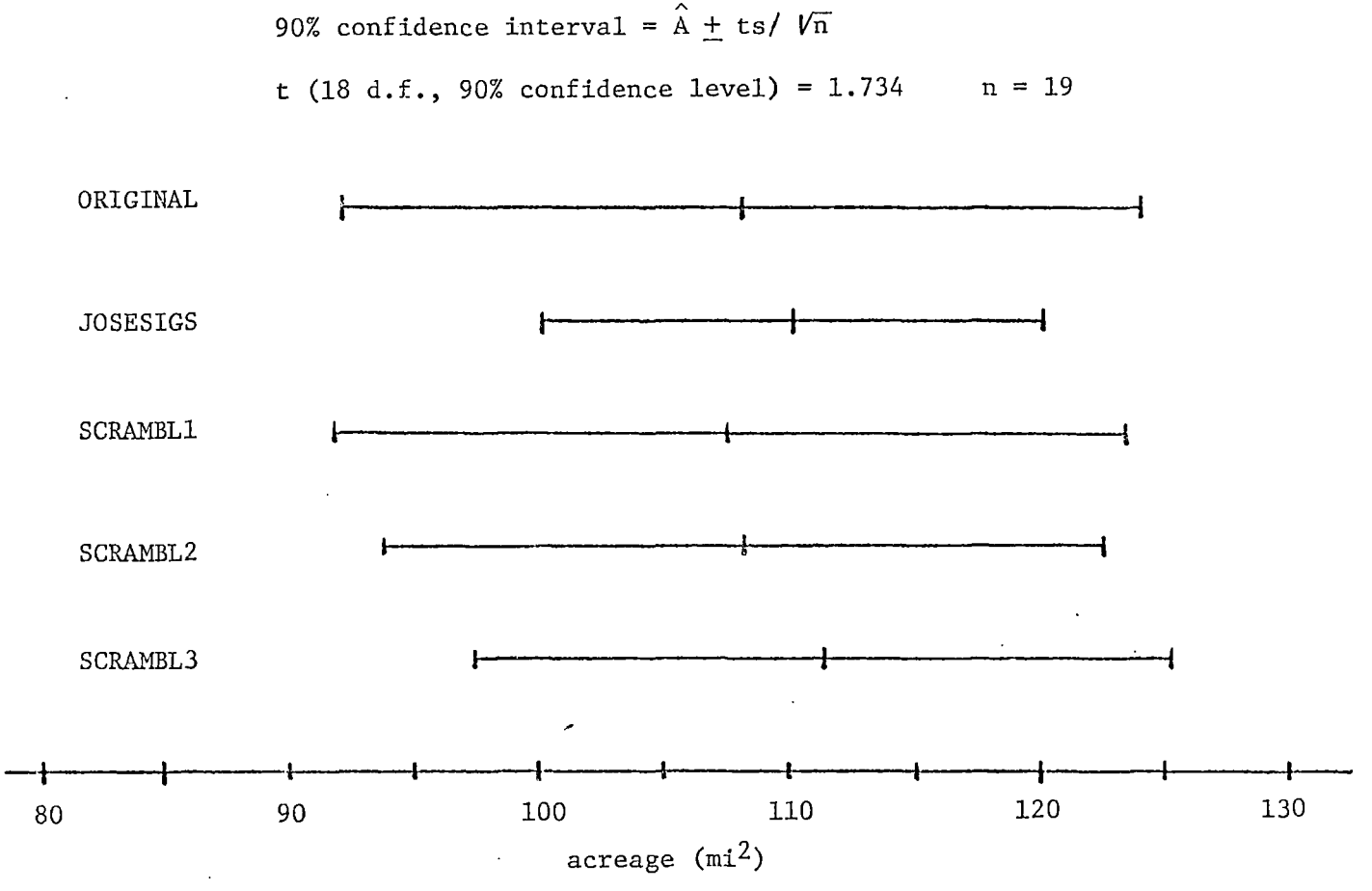


Figure 20. 90% confidence intervals for acreage estimates - Experiment 5.

6) H_0: $\mu_2 = \mu_4$, vs. H_1: $\mu_2 \neq \mu_4$

7) H_0: $\mu_2 = \mu_5$, vs. H_1: $\mu_2 \neq \mu_5$

8) H_0: $\mu_3 = \mu_4$, vs. H_1: $\mu_3 \neq \mu_4$

9) H_0: $\mu_3 = \mu_5$, vs. H_1: $\mu_3 \neq \mu_5$

10) H_0: $\mu_4 = \mu_5$, vs. $_1$: $\mu_4 \neq \mu_5$

The test statistic used was:

$$t_{n_1 + n_2 - 2} = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{1/n_1 + 1/n_2}}$$

where:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

To test at the 95% confidence level the tabular t value used is:

$$t_{36,.975} = 2.028$$

The calculations for the test of hypothesis one are given as an example:

$$n_1 = n_2 = 19$$

$$s_1 = 40.0, \; s_2 = 25.1$$

$$s_p^2 = \frac{(19-1)\ (40.0)^2 + (19-1)(25.1)^2}{19 + 19 - 2} = 1115.0$$

$$s_p = 33.4$$

$$t_{36} = \frac{108.1 - 110.1}{33.4 \sqrt{1/19 + 1/19}} = -0.185$$

Since $-t_{36,.975} < t_{36} < t_{36,.975}$ (i.e., $-2.034 < -0.185 < 2.034$) we accept the hypothesis H_0: $\mu_1 = \mu_2$. The results of the calculations for the other nine tests are shown in Table 19. It is seen that the

Table 19. Pairwise t test between acreage estimates.

Critical value at 95% confidence level with 36 degrees of freedom = 2.028

Test	H_0	H_1	t_{36}	accept or reject H_0
1	$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	-0.185	accept
2	$\mu_1 = \mu_3$	$\mu_1 \neq \mu_3$	0.037	accept
3	$\mu_1 = \mu_4$	$\mu_1 \neq \mu_4$	-0.006	accept
4	$\mu_1 = \mu_5$	$\mu_1 \neq \mu_5$	-0.266	accept
5	$\mu_2 = \mu_3$	$\mu_2 \neq \mu_3$	0.233	accept
6	$\mu_2 = \mu_4$	$\mu_2 \neq \mu_4$	0.192	accept
7	$\mu_2 = \mu_5$	$\mu_2 \neq \mu_5$	-0.123	accept
8	$\mu_3 = \mu_4$	$\mu_3 \neq \mu_4$	-0.045	accept
9	$\mu_3 = \mu_5$	$\mu_3 \neq \mu_5$	-0.307	accept
10	$\mu_4 = \mu_5$	$\mu_4 \neq \mu_5$	-0.274	accept

null hypothesis, that the acreage estimates are equal at the 95% significance level, is accepted in every case. A discussion of these results is given in the next chapter.

5.0 CONCLUSIONS

The results of this study demonstrate that acreage estimates and the variance of the estimates are not significantly different when different Landsat classifications of varying relative accuracy are used to obtain selection probabilities. In fact, there was a greater difference between different samples of the same classification than between the same samples (selecting the same sample units) using different classifications.

Inferences from the results of Experiment 1 can be made by viewing the plot of means and 90% confidence intervals (Figure 15). It is seen that the acreage estimates are in good agreement (ranging from 99 to 113) as are the 90% confidence intervals (ranging from 21 to 28). As was seen in subsequent experiments, however, there can be much greater differences between samples. The immediate conclusion of this result was that the sample size should be increased. Therefore, conclusive inferences about the effects of using different classifications were made only after further sampling was conducted.

Experiment 2 shows the effect that can be realized when different sample units are selected for the samples. Although the acreage estimates produced in this experiment were very close to those of Experiment 1, the variance of the estimate obtained using the classification SCRAMBL3 was several times larger than those of the other classifications (sample variance was 3060 for SCRAMBL3 and 221 for ORIGINAL). This was caused by the fact that those sample units of the SCRAMBL3 sample not in common with the samples of the other classifications (sample units 67 and 86) gave acreage estimates, A_i, different from

the others. This is caused by poor correlation between the Landsat data and the ground truth data for these sample units. The acreage estimate obtained from the entire SCRAMBL3 sample was close to the others, however, just because the very large acreage estimate of sample unit 67 was compensated for by the very small estimate of sample unit 86.

The effect of greatly differing sample variances (and also greatly differing acreage estimates, which is equally as possible but was not realized in this experiment) is a very significant effect caused by the specific classification used to obtain selection probabilities. This effect would be minimized, however, by taking a sample of sufficient size. If the sample of Experiment 2 using the SCRAMBL3 classification were used to compute the sample size necessary for a 10% allowable error we would get:

$$n' = \frac{t^2 \, C.V.^2}{A.E.^2} = \frac{(2.132)^2 \, (.226)^2}{(.10)^2} = 23.21$$

$$n = \frac{n'}{1 + \frac{n'}{N}} = \frac{23.21}{1 + \frac{23.21}{100}} = 18.84 \simeq 19$$

If a sample of this size were taken, the effects of those sample units showing poor correlation between Landsat data and ground truth would be minimized.

The only conclusion drawn from the results of Experiment 3 was that sampling is not an effective tool for estimating the acreage of this particular ground cover class. Except for small, isolated mountain meadows, it occurs primarily as stringers along rivers and streams. This is in contrast to the conifer class which had much broader patterns of distribution generally related to elevation, slope, and aspect.

The training field classification accuracy for meadow was worse than that of conifer (73.2% vs. 95.5%) but this was not believed to be responsible for this result. If this were the case, then the classifications where test set accuracy of the conifer class was reduced would have caused the same result. This was not observed.

The results of Experiment 4 show that there is a much larger difference between the results of individual samples than between the results obtained using different classifications. The only conclusion drawn from this observation was that sample size should be increased (Dixon, 1978).

It was thus decided to perform Experiment 5. It was hoped that with a sample of sufficient size, within sample variance could be reduced and more meaningful inferences made about differences between classifications. The same sample of the 19 sample units for which ground truth measurements were known was performed using each of the 5 classifications to obtain selection probabilities. A simple pairwise t test for the equality of means was performed to make inferences about the results.

In each of the 10 tests for equality of acreage estimates between 2 classifications, it was seen that there was no difference at the 95% significance level. Differences were not significant in each case even down to the 70% significance level. It was thus concluded that there was no significant difference in the acreage estimates obtained by using five different classifications to obtain selection probabilities.

An interesting observation which can be made from the results of Experiment 5 is that of the calculated sample size for an allowable

error of 10% and tabular t value for 90% confidence level. The calculations for necessary sample size using the data for the sample using the ORIGINAL classification give a value of three sample units required. The same calculations specifying an allowable error of 5% and 95% confidence level give a required sample size of 12 sample units. The experiments performed here have shown that a sample of size five is inadequate to make the desired inferences because of excessively large sample variances in some cases. Thus, it was concluded that specifying an allowable error of 10% and confidence level of 90% was not conservative enough to provide a sample of adequate size for these purposes. It may be wise to be fairly conservative in calculation of required sample size (in this case by specifying an allowable error of 5% instead of 10% and confidence level of 95% instead of 90%).

Because of the wide degree of variability in sample variance noted among the several samples made, no attempt was made to test the significance of the differences between the sample variances observed. It is assumed that variance would be fairly closely related to the degree of correlation between Landsat and ground truth data. As has been seen, however, sample variance can be somewhat controlled, or at least compensated for, by selecting a sample of adequate size.

One aspect of this study which was believed to hinder the results was inaccuracy of the ground truth data, due to reasons discussed in Section 4.3. This factor introduced a possible variability due to differences in ground truth measurements and actual ground truth. This may have tended to obscure the variability between Landsat predictions which was of primary concern to the study. However, since the effect is probably the same for all Landsat classifications, it

is not considered a great enough factor to invalidate the conclusions made. This problem could have been minimized if more up-to-date topographic maps at a scale of 1:24,000 were available for use as a ground truth map.

The original hypothesis, to be proven or disproven by this study is: relative classification accuracy has a significant effect on final acreage estimates obtained when using Landsat data to obtain selection probabilities in a variable probability sampling scheme. I believe that the results of this study are sufficient to disprove this hypothesis. Samples were performed using Landsat classifications of varying relative accuracy and differences were found to be insignificant at the 95% significance level, thus supporting subjective analysis of the results.

LITERATURE CITED

Anuta, Paul E. 1973. Geometric correction of ERTS-1 digital multispectral scanner data. LARS Information Note: 103073. Purdue Univ., West Lafayette, Ind. 23 pp.

Bowden, David C. 1978. Personal communication, April 27, 1978. Colorado State Univ., Fort Collins, Co.

Cochran, William G. 1963. Sampling techniques. John Wiley and Sons, Inc., New York, N.Y. 413 pp.

Davis, John C. 1973. Statistics and data analysis in geology. John Wiley and Sons, Inc., New York, N.Y. 550 pp.

Dixon, Gary E. 1978. Personal communication. Dr. Gary Dixon provided assistance on numerous occasions throughout the study. Rocky Mountain For. and Range Exp. Sta., Fort Collins, Co.

Driscoll, Richard S., and Richard E. Francis. 1970. Multistage, multiseasonal, and multiband imagery to identify and qualify non-forest vegetation resources. Rocky Mountain For. and Range Exp. Sta., Fort Collins, Co. 65 pp.

Fleming, M.D., J.S. Berkebile, and R.M. Hoffer. 1975. Computer aided analysis of Landsat-1 MSS data: a comparison of three approaches including a "modified clustering" approach. LARS Information Note: 072475. Purdue Univ., West Lafayette, Ind. 9 pp.

Guenther, William C. 1965. Concepts of statistical inference. McGraw-Hill Book Company, New York, N.Y. 353 pp.

Hansen, M.H., and W.N. Hurwitz. 1943. On the theory of sampling from finite populations. Annals of Math. Stat. Vol. XIV, No. 4, pp. 333-362.

Heller, R.C., and J.F. Wear. 1969. Sampling forest insect epidemics with color films. Proc. 6th Intl. Symp. on Remote Sensing of Env., Univ. of Michigan, Ann Arbor, Mich. pp. 1157-1167.

Husch, Bertram, Charles I. Miller, and Thomas W. Beers. 1972. Forest mensuration. The Ronald Press Company, New York, N.Y. 410 pp.

Langley, Philip G. 1969. New multistage sampling techniques using space and aircraft imagery for forest inventory. Proc. 6th Intl. Symp. on Remote Sensing of Env. Univ. of Michigan, Ann Arbor, Mich. pp. 1179-1192.

Langley, Philp G., Robert C. Aldrich, and Robert C. Heller. 1969. Multistage sampling of the forest resource using space photography - an Apollo 9 case study. Second Annual Earth Resources Aircraft Program Status Review Proceedings. pp 19-1 - 19-21.

Langley, Philip G. 1971. Multistage sampling of earth resources with aerial and space photography. In Monitoring Earth Resources from Aircraft and Spacecraft, Robert N. Colwell (ed.), NASA SP-275, pp. 129-141.

Maxwell, Eugene L. 1976. Multivariate system analysis of multispectral imagery. Photogrammetric Engineering and Remote Sensing. Vol. XLII, No. 9, pp. 1173-1186.

Minter, Ruth T. 1972. Computer program documentation, ISOCLS, iterative self-organizing clustering program. NASA, Manned Spacecraft Center, Houston, Texas.

Nichols, J.D., D.T. Lauer, M. Gialdini, and W. Orme. 1973. ERTS as an aid in timber volume inventory. A special report to Office of Applications, NASA Hdqtrs. Univ. of California, Berkeley, Cal. 31 pp.

Pernia, José E. 1978. Hydrologic parameters from Landsat imagery for Williams Fork watershed. M.S. Thesis, Colorado State Univ., Fort Collins, Co. 124 pp.

Ranson, K. Jon, and Frank Itkowsky. 1977. Middle Park inventory analysis. Final Report, U.S. Forest Service Cooperative Agreement 16-630-CA. Colorado State Univ., Fort Collins, Co. 51 pp.

Rhode, Wayne G. 1976. Calculation of optimum size of primary sample units (PSU). Unpublished paper. 11 pp.

Rhode, Wayne G., and Gregory R. Johnson. 1976. Application of PPS sampling to estimating acreage of forest land with satellite and aircraft data. Unpublished paper. 15 pp.

Smith, James A., Lee D. Miller, and Tom Ells. 1972. Pattern recognition routines for graduate training in the automatic analysis of remote sensing imagery - RECOG. Science Series No. 34, Colorado State Univ., Fort Collins, Co.

Titus, S., M. Gialdini, and J. Nichols. 1975. A total timber resource inventory based upon manual and automated analysis of LANDSAT-1 and supporting aircraft data using stratified multistage sampling techniques. Proc. 10th Intl. Symp. on Remote Sensing of Env., Univ. of Michigan, Ann Arbor, Mich. pp. 1093-1099.

APPENDIX A

Confusion Matrices Showing Classification Accuracy

Table 20. Training field classification accuracy - ORIGINAL.

Points Classified As:

Training Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	62	2		3				
Conifer	8	121					1	
Mt. Grass			17				1	
Meadow	28		1	93	6			
Mt. Shrub					10		2	
Water						129	1	
Sage			7		4		169	6

Percentage Classified As:

Training Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	92.5	3.0		4.5				
Conifer	6.1	93.1					0.8	
Mt. Grass			94.4				5.6	
Meadow	21.9		0.8	72.7	4.7			
Mt. Shrub					83.3		16.7	
Water						99.2	0.8	
Sage			3.8		2.2		90.9	3.2

Percent of points correctly classified = 89.6

Table 21. Test field classification accuracy - JOSESIGS.

Points Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	17	28		16	6			
Conifer	2	127						1
Mt. Grass			2					16
Meadow			2	122	4			
Mt. Shrub			1	6	3			2
Water						127		3
Sage*							-	

Percentage Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	25.4	41.8		23.9	9.0			
Conifer	1.5	97.7						0.8
Mt. Grass			11.1					88.9
Meadow			1.6	95.3	3.1			
Mt. Shrub			8.3	50.0	25.0			16.7
Water						97.7		2.3
Sage*							-	

Percent of Points Correctly Classified = 82.1

* A SAGE class was not defined for this classification

Table 22. Test field classification accuracy - SCRAMBL1.

Points Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	60	2		5				
Conifer	6	115			1		8	
Mt. Grass			16		1		1	
Meadow	38		1	76	11		2	
Mt. Shrub					7		5	
Water						129	1	
Sage			8		7		166	5

Percentage Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	89.6	3.0		7.5				
Conifer	4.6	88.5			0.8		6.2	
Mt. Grass			88.9		5.6		5.6	
Meadow	29.7		0.8	59.4	8.6		1.6	
Mt. Shrub					58.3		41.7	
Water						99.2	0.8	
Sage			4.3		3.8		89.2	2.7

Percent of Points Correctly Classified = 84.8

Table 23. Test field classification accuracy - SCRAMBL2.

Points Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	49	4		9	1		4	
Conifer	17	93	1			2	17	
Mt. Grass			14		1		3	
Meadow	39	4	1	63	17		4	
Mt. Shrub					7		5	
Water	5					124	1	
Sage	3		16	1	17		145	4

Percentage Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	73.1	6.0		13.4	1.5		6.0	
Conifer	13.1	71.5	0.8			1.5	13.1	
Mt. Grass			77.8		5.6		16.7	
Meadow	30.5	3.1	0.8	49.2	13.3		3.1	
Mt. Shrub					58.3		41.7	
Water	3.8					95.4	0.8	
Sage	1.6		8.6	0.5	9.1		78.0	2.2

Percent of Points Correctly Classified = 73.8

Table 24. Test field classification accuracy - SCRAMBL3.

Points Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	45	16		4	2			
Conifer	7	122	1					
Mt. Grass			15	1			2	
Meadow	29	6		83	9		1	
Mt. Shrub				1	7		4	
Water	7					122	1	
Sage	1		15	9	16		140	5

Percentage Classified As:

Test Field	Decid.	Conif.	Grass	Meadow	Shrub	Water	Sage	Distb.
Deciduous	67.2	23.9		6.0	3.0			
Conifer	5.4	93.8	0.8					
Mt. Grass			83.3	5.6			11.1	
Meadow	22.7	4.7		64.8	7.0		0.8	
Mt. Shrub				8.3	58.3		33.3	
Water	5.4					93.8	0.8	
Sage	0.5		8.1	4.8	8.6		75.3	2.7

Percent of Points Correctly Classified = 79.6

APPENDIX B

Sample Lists

Table 25. Sample list for CONIFER sample using ORIGINAL classification.

PRIMARY SAMPLE UNITS FOR CONIF								
PSU #	LINES			COLS.			X(I)	SUM X(I)
1	1	20	1	1	20	1	0.0	0.000
2	1	20	1	21	40	1	30.0	30.000
3	1	20	1	41	60	1	11.0	41.000
4	1	20	1	61	80	1	0.0	41.000
5	1	20	1	81	100	1	0.0	41.000
6	1	20	1	101	120	1	4.0	45.000
7	1	20	1	121	140	1	94.0	139.000
8	1	20	1	141	160	1	70.0	209.000
9	1	20	1	161	180	1	143.0	352.000
10	1	20	1	181	200	1	116.0	468.000
11	21	40	1	1	20	1	4.0	472.000
12	21	40	1	21	40	1	0.0	472.000
13	21	40	1	41	60	1	0.0	472.000
14	21	40	1	61	80	1	0.0	472.000
15	21	40	1	81	100	1	0.0	472.000
16	21	40	1	101	120	1	0.0	472.000
17	21	40	1	121	140	1	0.0	472.000
18	21	40	1	141	160	1	5.0	477.000
19	21	40	1	161	180	1	106.0	583.000
20	21	40	1	181	200	1	24.0	607.000
21	41	60	1	1	20	1	6.0	613.000
22	41	60	1	21	40	1	1.0	614.000
23	41	60	1	41	60	1	1.0	615.000
24	41	60	1	61	80	1	0.0	615.000
25	41	60	1	81	100	1	1.0	616.000
26	41	60	1	101	120	1	5.0	621.000
27	41	60	1	121	140	1	5.0	626.000
28	41	60	1	141	160	1	6.0	632.000
29	41	60	1	161	180	1	16.0	648.000
30	41	60	1	181	200	1	53.0	701.000
31	61	80	1	1	20	1	35.0	736.000
32	61	80	1	21	40	1	0.0	736.000
33	61	80	1	41	60	1	6.0	742.000
34	61	80	1	61	80	1	3.0	745.000
35	61	80	1	81	100	1	0.0	745.000
36	61	80	1	101	120	1	0.0	745.000
37	61	80	1	121	140	1	0.0	745.000
38	61	80	1	141	160	1	6.0	751.000
39	61	80	1	161	180	1	0.0	751.000
40	61	80	1	181	200	1	43.0	794.000
41	81	100	1	1	20	1	17.0	811.000
42	81	100	1	21	40	1	2.0	813.000
43	81	100	1	41	60	1	4.0	817.000
44	81	100	1	61	80	1	11.0	828.000
45	81	100	1	81	100	1	18.0	846.000
46	81	100	1	101	120	1	25.0	871.000
47	81	100	1	121	140	1	6.0	877.000
48	81	100	1	141	160	1	1.0	878.000
49	81	100	1	161	180	1	14.0	892.000
50	81	100	1	181	200	1	57.0	949.000
51	101	120	1	1	20	1	95.0	1044.000
52	101	120	1	21	40	1	12.0	1056.000
53	101	120	1	41	60	1	1.0	1057.000
54	101	120	1	61	80	1	0.0	1057.000
55	101	120	1	81	100	1	83.0	1140.000
56	101	120	1	101	120	1	147.0	1287.000
57	101	120	1	121	140	1	15.0	1302.000
58	101	120	1	141	160	1	5.0	1307.000
59	101	120	1	161	180	1	40.0	1347.000
60	101	120	1	181	200	1	153.0	1500.000
61	121	140	1	1	20	1	73.0	1573.000
62	121	140	1	21	40	1	93.0	1666.000
63	121	140	1	41	60	1	3.0	1669.000
64	121	140	1	61	80	1	18.0	1687.000
65	121	140	1	81	100	1	8.0	1695.000
66	121	140	1	101	120	1	137.0	1832.000
67	121	140	1	121	140	1	91.0	1923.000
68	121	140	1	141	160	1	15.0	1938.000
69	121	140	1	161	180	1	85.0	2023.000
70	121	140	1	181	200	1	174.0	2197.000

Table 25. (Continued)

PRIMARY SAMPLE UNITS FOR CONIF PSU =		LINES			COLS.		X(I)	SUM X(I)
71	141	160	1	1	20	1	191.0	2378.000
72	141	160	1	21	40	1	159.0	2537.000
73	141	160	1	41	60	1	64.0	2601.000
74	141	160	1	61	80	1	29.0	2630.000
75	141	160	1	81	100	1	2.0	2632.000
76	141	160	1	101	120	1	3.0	2635.000
77	141	160	1	121	140	1	190.0	2825.000
78	141	160	1	141	160	1	190.0	3015.000
79	141	160	1	161	180	1	143.0	3158.000
80	141	160	1	181	200	1	172.0	3330.000
81	161	180	1	1	20	1	121.0	3451.000
82	161	180	1	21	40	1	113.0	3564.000
83	161	180	1	41	60	1	190.0	3754.000
84	161	180	1	61	80	1	134.0	3888.000
85	161	180	1	81	100	1	27.0	3915.000
86	161	180	1	101	120	1	22.0	3937.000
87	161	180	1	121	140	1	36.0	3973.000
88	161	180	1	141	160	1	146.0	4119.000
89	161	180	1	161	180	1	236.0	4355.000
90	161	180	1	181	200	1	194.0	4549.000
91	181	200	1	1	20	1	89.0	4638.000
92	181	200	1	21	40	1	60.0	4698.000
93	181	200	1	41	60	1	177.0	4875.000
94	181	200	1	61	80	1	242.0	5117.000
95	181	200	1	81	100	1	213.0	5330.000
96	181	200	1	101	120	1	119.0	5449.000
97	181	200	1	121	140	1	110.0	5559.000
98	181	200	1	141	160	1	10.0	5569.000
99	181	200	1	161	180	1	40.0	5609.000
100	181	200	1	181	200	1	165.0	5774.000

MEAN = 57.74000
STANDARD DEVIATION = 66.62588
COEFFICIENT OF VARIATION = 1.18883

Table 26. Sample list for CONIFER sample using JOSESIGS classification.

PRIMARY SAMPLE UNITS FOR PSU =		LINES	AAAA1		BBBB2 COLS.	CCCC3	X(I)	SUM X(I)
1	1	20	1	1	20	1	0.0	0.000
2	1	20	1	21	40	1	31.0	31.000
3	1	20	1	41	60	1	15.0	46.000
4	1	20	1	61	80	1	0.0	46.000
5	1	20	1	81	100	1	0.0	46.000
6	1	20	1	101	120	1	7.0	53.000
7	1	20	1	121	140	1	140.0	193.000
8	1	20	1	141	160	1	124.0	317.000
9	1	20	1	161	180	1	232.0	549.000
10	1	20	1	181	200	1	187.0	736.000
11	21	40	1	1	20	1	3.0	739.000
12	21	40	1	21	40	1	1.0	740.000
13	21	40	1	41	60	1	0.0	740.000
14	21	40	1	61	80	1	0.0	740.000
15	21	40	1	81	100	1	0.0	740.000
16	21	40	1	101	120	1	0.0	740.000
17	21	40	1	121	140	1	0.0	740.000
18	21	40	1	141	160	1	6.0	746.000
19	21	40	1	161	180	1	118.0	864.000
20	21	40	1	181	200	1	34.0	898.000
21	41	60	1	1	20	1	9.0	907.000
22	41	60	1	21	40	1	2.0	909.000
23	41	60	1	41	60	1	0.0	909.000
24	41	60	1	61	80	1	1.0	910.000
25	41	60	1	81	100	1	1.0	911.000
26	41	60	1	101	120	1	5.0	916.000
27	41	60	1	121	140	1	7.0	923.000
28	41	60	1	141	160	1	6.0	929.000
29	41	60	1	161	180	1	22.0	951.000
30	41	60	1	181	200	1	69.0	1020.000
31	61	80	1	1	20	1	67.0	1087.000
32	61	80	1	21	40	1	1.0	1088.000
33	61	80	1	41	60	1	7.0	1095.000
34	61	80	1	61	80	1	3.0	1098.000
35	61	80	1	81	100	1	0.0	1098.000
36	61	80	1	101	120	1	0.0	1098.000
37	61	80	1	121	140	1	0.0	1098.000
38	61	80	1	141	160	1	3.0	1101.000
39	61	80	1	161	180	1	0.0	1101.000
40	61	80	1	181	200	1	55.0	1156.000
41	81	100	1	1	20	1	76.0	1232.000
42	81	100	1	21	40	1	5.0	1237.000
43	81	100	1	41	60	1	2.0	1239.000
44	81	100	1	61	80	1	18.0	1257.000
45	81	100	1	81	100	1	30.0	1287.000
46	81	100	1	101	120	1	49.0	1336.000
47	81	100	1	121	140	1	7.0	1343.000
48	81	100	1	141	160	1	0.0	1343.000
49	81	100	1	161	180	1	19.0	1362.000
50	81	100	1	181	200	1	111.0	1473.000
51	101	120	1	1	20	1	185.0	1658.000
52	101	120	1	21	40	1	35.0	1703.000
53	101	120	1	41	60	1	1.0	1704.000
54	101	120	1	61	80	1	0.0	1704.000
55	101	120	1	81	100	1	124.0	1828.000
56	101	120	1	101	120	1	261.0	2089.000
57	101	120	1	121	140	1	54.0	2143.000
58	101	120	1	141	160	1	8.0	2151.000
59	101	120	1	161	180	1	48.0	2199.000
60	101	120	1	181	200	1	230.0	2429.000
61	121	140	1	1	20	1	143.0	2572.000
62	121	140	1	21	40	1	158.0	2730.000
63	121	140	1	41	60	1	10.0	2740.000
64	121	140	1	61	80	1	16.0	2756.000
65	121	140	1	81	100	1	22.0	2778.000
66	121	140	1	101	120	1	175.0	2953.000
67	121	140	1	121	140	1	211.0	3164.000
68	121	140	1	141	160	1	96.0	3260.000
69	121	140	1	161	180	1	118.0	3378.000
70	121	140	1	181	200	1	225.0	3603.000

Table 26. (Continued)

PRIMARY SAMPLE UNITS FOR PSU #	LINES		AAAA1	COLS.		BBBB2	X(I)	CCCC3 SUM X(I)
71	141	160	1	1	20	1	221.0	3934.000
72	141	160	1	21	40	1	251.0	4065.000
73	141	160	1	41	60	1	111.0	4196.000
74	141	160	1	61	80	1	37.0	4233.000
75	141	160	1	81	100	1	3.0	4236.000
76	141	160	1	101	120	1	6.0	4242.000
77	141	160	1	121	140	1	239.0	4481.000
78	141	160	1	141	160	1	277.0	4758.000
79	141	160	1	141	180	1	286.0	5044.000
80	141	160	1	181	200	1	253.0	5297.000
81	161	180	1	1	20	1	178.0	5475.000
82	161	180	1	21	40	1	205.0	5680.000
83	161	180	1	41	60	1	240.0	5920.000
84	161	180	1	61	80	1	183.0	6103.000
85	161	180	1	81	100	1	72.0	6175.000
86	161	180	1	101	120	1	43.0	6218.000
87	161	180	1	121	140	1	46.0	6264.000
88	161	180	1	141	160	1	182.0	6446.000
89	161	180	1	161	180	1	288.0	6734.000
90	161	180	1	181	200	1	273.0	7007.000
91	181	200	1	1	20	1	137.0	7144.000
92	181	200	1	21	40	1	122.0	7266.000
93	181	200	1	41	60	1	246.0	7512.000
94	181	200	1	61	80	1	270.0	7782.000
95	181	200	1	81	100	1	280.0	8062.000
96	181	200	1	101	120	1	205.0	8267.000
97	181	200	1	121	140	1	178.0	8445.000
98	181	200	1	141	160	1	14.0	8459.000
99	181	200	1	141	180	1	46.0	8505.000
100	181	200	1	181	200	1	203.0	8708.000

MEAN = 87.08000
STANDARD DEVIATION = 96.99824
COEFFICIENT OF VARIATION = 1.11309

Table 27. Sample list for CONIFER sample using SCRAMBL1 classification.

PRIMARY SAMPLE UNITS FOR CONIF

PSU #		LINES			COLS.		Y(I)	SUM X(I)
1	1	20	1	1	20	1	0.0	0.000
2	1	20	1	21	40	1	26.0	26.000
3	1	20	1	41	60	1	11.0	37.000
4	1	20	1	61	80	1	0.0	37.000
5	1	20	1	81	100	1	0.0	37.000
6	1	20	1	101	120	1	4.0	41.000
7	1	20	1	121	140	1	90.0	131.000
8	1	20	1	141	160	1	69.0	200.000
9	1	20	1	161	180	1	139.0	339.000
10	1	20	1	181	200	1	110.0	449.000
11	21	40	1	1	20	1	4.0	453.000
12	21	40	1	21	40	1	0.0	453.000
13	21	40	1	41	60	1	0.0	453.000
14	21	40	1	61	80	1	0.0	453.000
15	21	40	1	81	100	1	0.0	453.000
16	21	40	1	101	120	1	0.0	453.000
17	21	40	1	121	140	1	0.0	453.000
18	21	40	1	141	160	1	4.0	457.000
19	21	40	1	161	180	1	102.0	559.000
20	21	40	1	181	200	1	23.0	582.000
21	41	60	1	1	20	1	6.0	588.000
22	41	60	1	21	40	1	1.0	589.000
23	41	60	1	41	60	1	1.0	590.000
24	41	60	1	61	80	1	0.0	590.000
25	41	60	1	81	100	1	1.0	591.000
26	41	60	1	101	120	1	5.0	596.000
27	41	60	1	121	140	1	5.0	601.000
28	41	60	1	141	160	1	5.0	606.000
29	41	60	1	161	180	1	15.0	621.000
30	41	60	1	181	200	1	53.0	674.000
31	61	80	1	1	20	1	34.0	708.000
32	61	80	1	21	40	1	0.0	708.000
33	61	80	1	41	60	1	5.0	713.000
34	61	80	1	61	80	1	2.0	715.000
35	61	80	1	81	100	1	0.0	715.000
36	61	80	1	101	120	1	0.0	715.000
37	61	80	1	121	140	1	0.0	715.000
38	61	80	1	141	160	1	6.0	721.000
39	61	80	1	161	180	1	0.0	721.000
40	61	80	1	181	200	1	41.0	762.000
41	81	100	1	1	20	1	17.0	779.000
42	81	100	1	21	40	1	2.0	781.000
43	81	100	1	41	60	1	4.0	785.000
44	81	100	1	61	80	1	10.0	795.000
45	81	100	1	81	100	1	17.0	812.000
46	81	100	1	101	120	1	24.0	836.000
47	81	100	1	121	140	1	6.0	842.000
48	81	100	1	141	160	1	1.0	843.000
49	81	100	1	161	180	1	13.0	856.000
50	81	100	1	181	200	1	54.0	910.000
51	101	120	1	1	20	1	90.0	1000.000
52	101	120	1	21	40	1	11.0	1011.000
53	101	120	1	41	60	1	1.0	1012.000
54	101	120	1	61	80	1	0.0	1012.000
55	101	120	1	81	100	1	80.0	1092.000
56	101	120	1	101	120	1	140.0	1232.000
57	101	120	1	121	140	1	14.0	1246.000
58	101	120	1	141	160	1	4.0	1250.000
59	101	120	1	161	180	1	38.0	1288.000
60	101	120	1	181	200	1	149.0	1437.000
61	121	140	1	1	20	1	69.0	1506.000
62	121	140	1	21	40	1	90.0	1596.000
63	121	140	1	41	60	1	3.0	1599.000
64	121	140	1	61	80	1	17.0	1616.000
65	121	140	1	81	100	1	8.0	1624.000
66	121	140	1	101	120	1	125.0	1749.000
67	121	140	1	121	140	1	88.0	1837.000
68	121	140	1	141	160	1	15.0	1852.000
69	121	140	1	161	180	1	82.0	1934.000
70	121	140	1	181	200	1	165.0	2099.000

Table 27. (Continued)

PRIMARY SAMPLE UNITS FOR CONIF

PSU #	LINES			COLS.			X(I)	SUM X(I)
71	141	160	1	1	20	1	172.0	2271.000
72	141	160	1	21	40	1	147.0	2418.000
73	141	160	1	41	60	1	43.0	2461.000
74	141	160	1	61	80	1	46.0	2507.000
75	141	160	1	81	100	1	2.0	2509.000
76	141	160	1	101	120	1	3.0	2512.000
77	141	160	1	121	140	1	188.0	2700.000
78	141	160	1	141	160	1	180.0	2880.000
79	141	160	1	161	180	1	134.0	3014.000
80	141	160	1	181	200	1	167.0	3181.000
81	161	180	1	1	20	1	114.0	3295.000
82	161	180	1	21	40	1	111.0	3406.000
83	161	180	1	41	60	1	178.0	3584.000
84	161	180	1	61	80	1	127.0	3711.000
85	161	180	1	81	100	1	25.0	3736.000
86	161	180	1	101	120	1	22.0	3758.000
87	161	180	1	121	140	1	34.0	3792.000
88	161	180	1	141	160	1	140.0	3932.000
89	161	180	1	161	180	1	223.0	4155.000
90	161	180	1	181	200	1	187.0	4342.000
91	181	200	1	1	20	1	87.0	4429.000
92	181	200	1	21	40	1	58.0	4487.000
93	181	200	1	41	60	1	168.0	4655.000
94	181	200	1	61	80	1	232.0	4887.000
95	181	200	1	81	100	1	202.0	5089.000
96	181	200	1	101	120	1	112.0	5201.000
97	181	200	1	121	140	1	107.0	5308.000
98	181	200	1	141	160	1	9.0	5317.000
99	181	200	1	161	180	1	37.0	5354.000
100	181	200	1	181	200	1	155.0	5509.000

MEAN = 55.09000
STANDARD DEVIATION = 65.46232
COEFFICIENT OF VARIATION = 1.18828

Table 28. Sample list for CONIFER sample using SCRAMBL2 classification.

PRIMARY SAMPLE UNITS FOR CONIF

PSU #	LINES			COLS.			X(I)	SUM X(I)
1	1	20	1	1	20	1	0.0	0.000
2	1	20	1	21	40	1	19.0	19.000
3	1	20	1	41	60	1	10.0	29.000
4	1	20	1	61	80	1	0.0	29.000
5	1	20	1	81	100	1	1.0	30.000
6	1	20	1	101	120	1	5.0	35.000
7	1	20	1	121	140	1	81.0	116.000
8	1	20	1	141	160	1	45.0	161.000
9	1	20	1	161	180	1	130.0	291.000
10	1	20	1	181	200	1	98.0	409.000
11	21	40	1	1	20	1	3.0	412.000
12	21	40	1	21	40	1	0.0	412.000
13	21	40	1	41	60	1	0.0	412.000
14	21	40	1	61	80	1	0.0	412.000
15	21	40	1	81	100	1	0.0	412.000
16	21	40	1	101	120	1	0.0	412.000
17	21	40	1	121	140	1	0.0	412.000
18	21	40	1	141	160	1	4.0	416.000
19	21	40	1	161	180	1	88.0	504.000
20	21	40	1	181	200	1	23.0	527.000
21	41	60	1	1	20	1	7.0	534.000
22	41	60	1	21	40	1	3.0	537.000
23	41	60	1	41	60	1	1.0	538.000
24	41	60	1	61	80	1	0.0	538.000
25	41	60	1	81	100	1	2.0	540.000
26	41	60	1	101	120	1	4.0	544.000
27	41	60	1	121	140	1	5.0	549.000
28	41	60	1	141	160	1	4.0	553.000
29	41	60	1	161	180	1	14.0	567.000
30	41	60	1	181	200	1	53.0	620.000
31	61	80	1	1	20	1	32.0	652.000
32	61	80	1	21	40	1	0.0	652.000
33	61	80	1	41	60	1	6.0	658.000
34	61	80	1	61	80	1	2.0	660.000
35	61	80	1	81	100	1	1.0	661.000
36	61	80	1	101	120	1	1.0	662.000
37	61	80	1	121	140	1	6.0	668.000
38	61	80	1	141	160	1	9.0	677.000
39	61	80	1	161	180	1	0.0	677.000
40	61	80	1	181	200	1	36.0	713.000
41	81	100	1	1	20	1	21.0	734.000
42	81	100	1	21	40	1	2.0	736.000
43	81	100	1	41	60	1	7.0	743.000
44	81	100	1	61	80	1	10.0	753.000
45	81	100	1	81	100	1	16.0	769.000
46	81	100	1	101	120	1	23.0	792.000
47	81	100	1	121	140	1	5.0	797.000
48	81	100	1	141	160	1	6.0	803.000
49	81	100	1	161	180	1	12.0	815.000
50	81	100	1	181	200	1	54.0	869.000
51	101	120	1	1	20	1	88.0	957.000
52	101	120	1	21	40	1	11.0	968.000
53	101	120	1	41	60	1	1.0	969.000
54	101	120	1	61	80	1	0.0	969.000
55	101	120	1	81	100	1	69.0	1038.000
56	101	120	1	101	120	1	124.0	1162.000
57	101	120	1	121	140	1	22.0	1184.000
58	101	120	1	141	160	1	4.0	1188.000
59	101	120	1	161	180	1	37.0	1225.000
60	101	120	1	181	200	1	139.0	1364.000
61	121	140	1	1	20	1	71.0	1435.000
62	121	140	1	21	40	1	84.0	1519.000
63	121	140	1	41	60	1	3.0	1522.000
64	121	140	1	61	80	1	14.0	1536.000
65	121	140	1	81	100	1	9.0	1545.000
66	121	140	1	101	120	1	112.0	1657.000
67	121	140	1	121	140	1	88.0	1745.000
68	121	140	1	141	160	1	24.0	1769.000
69	121	140	1	161	180	1	76.0	1845.000
70	121	140	1	181	200	1	148.0	1993.000

Table 28. (Continued)

PRIMARY SAMPLE UNITS FOR CONIF PSU #	LINES			COLS.			X(I)	SUM X(I)
71	141	160	1	1	20	1	147.0	2140.000
72	141	160	1	21	40	1	132.0	2272.000
73	141	160	1	41	60	1	56.0	2328.000
74	141	160	1	61	80	1	24.0	2352.000
75	141	160	1	81	100	1	7.0	2359.000
76	141	160	1	101	120	1	5.0	2364.000
77	141	160	1	121	140	1	163.0	2527.000
78	141	160	1	141	160	1	163.0	2690.000
79	141	160	1	141	180	1	127.0	2817.000
80	141	160	1	181	200	1	139.0	2956.000
81	161	180	1	1	20	1	97.0	3053.000
82	161	180	1	21	40	1	101.0	3154.000
83	161	180	1	41	60	1	154.0	3308.000
84	161	180	1	61	80	1	117.0	3425.000
85	161	180	1	81	100	1	26.0	3451.000
86	161	180	1	101	120	1	25.0	3476.000
87	161	180	1	121	140	1	30.0	3506.000
88	161	180	1	141	160	1	120.0	3626.000
89	161	180	1	161	180	1	188.0	3814.000
90	161	180	1	181	200	1	169.0	3983.000
91	181	200	1	1	20	1	80.0	4063.000
92	181	200	1	21	40	1	60.0	4123.000
93	181	200	1	41	60	1	143.0	4266.000
94	181	200	1	61	80	1	202.0	4468.000
95	181	200	1	81	100	1	183.0	4651.000
96	181	200	1	101	120	1	106.0	4757.000
97	181	200	1	121	140	1	94.0	4851.000
98	181	200	1	141	160	1	13.0	4864.000
99	181	200	1	161	180	1	33.0	4897.000
100	181	200	1	181	200	1	135.0	5032.000

MEAN = 50.32000
STANDARD DEVIATION = 57.41144
COEFFICIENT OF VARIATION = 1.14093

Table 29. Sample list for CONIFER sample using SCRAMBL3 classification.

PRIMARY SAMPLE UNITS FOR CONIF								
PSU #	LINES			COLS.			X(I)	SUM X(I)
1	1	20	1	1	20	1	1.0	1.000
2	1	20	1	21	40	1	32.0	33.000
3	1	20	1	41	60	1	13.0	46.000
4	1	20	1	61	80	1	4.0	50.000
5	1	20	1	81	100	1	13.0	63.000
6	1	20	1	101	120	1	6.0	69.000
7	1	20	1	121	140	1	100.0	169.000
8	1	20	1	141	160	1	88.0	257.000
9	1	20	1	161	180	1	162.0	419.000
10	1	20	1	181	200	1	135.0	554.000
11	21	40	1	1	20	1	4.0	558.000
12	21	40	1	21	40	1	0.0	558.000
13	21	40	1	41	60	1	0.0	558.000
14	21	40	1	61	80	1	0.0	558.000
15	21	40	1	81	100	1	11.0	569.000
16	21	40	1	101	120	1	1.0	570.000
17	21	40	1	121	140	1	1.0	571.000
18	21	40	1	141	160	1	6.0	577.000
19	21	40	1	161	180	1	105.0	682.000
20	21	40	1	181	200	1	36.0	718.000
21	41	60	1	1	20	1	11.0	729.000
22	41	60	1	21	40	1	16.0	745.000
23	41	60	1	41	60	1	2.0	747.000
24	41	60	1	61	80	1	2.0	749.000
25	41	60	1	81	100	1	8.0	757.000
26	41	60	1	101	120	1	11.0	768.000
27	41	60	1	121	140	1	6.0	774.000
28	41	60	1	141	160	1	8.0	782.000
29	41	60	1	161	180	1	18.0	800.000
30	41	60	1	181	200	1	56.0	856.000
31	61	80	1	1	20	1	53.0	909.000
32	61	80	1	21	40	1	5.0	914.000
33	61	80	1	41	60	1	8.0	922.000
34	61	80	1	61	80	1	3.0	925.000
35	61	80	1	81	100	1	2.0	927.000
36	61	80	1	101	120	1	7.0	934.000
37	61	80	1	121	140	1	9.0	943.000
38	61	80	1	141	160	1	8.0	951.000
39	61	80	1	161	180	1	0.0	951.000
40	61	80	1	181	200	1	44.0	995.000
41	81	100	1	1	20	1	36.0	1031.000
42	81	100	1	21	40	1	8.0	1039.000
43	81	100	1	41	60	1	12.0	1051.000
44	81	100	1	61	80	1	12.0	1063.000
45	81	100	1	81	100	1	26.0	1089.000
46	81	100	1	101	120	1	36.0	1125.000
47	81	100	1	121	140	1	11.0	1136.000
48	81	100	1	141	160	1	8.0	1144.000
49	81	100	1	161	180	1	18.0	1162.000
50	81	100	1	181	200	1	80.0	1242.000
51	101	120	1	1	20	1	121.0	1363.000
52	101	120	1	21	40	1	25.0	1388.000
53	101	120	1	41	60	1	3.0	1391.000
54	101	120	1	61	80	1	0.0	1391.000
55	101	120	1	81	100	1	90.0	1481.000
56	101	120	1	101	120	1	166.0	1647.000
57	101	120	1	121	140	1	37.0	1684.000
58	101	120	1	141	160	1	10.0	1694.000
59	101	120	1	161	180	1	47.0	1741.000
60	101	120	1	181	200	1	170.0	1911.000
61	121	140	1	1	20	1	85.0	1996.000
62	121	140	1	21	40	1	106.0	2102.000
63	121	140	1	41	60	1	13.0	2115.000
64	121	140	1	61	80	1	20.0	2135.000
65	121	140	1	81	100	1	13.0	2148.000
66	121	140	1	101	120	1	139.0	2287.000
67	121	140	1	121	140	1	115.0	2402.000
68	121	140	1	141	160	1	48.0	2450.000
69	121	140	1	161	180	1	98.0	2548.000
70	121	140	1	181	200	1	175.0	2723.000

Table 29. (Continued)

PRIMARY SAMPLE UNITS FOR PSU =	LINES		CONIF	COLS.		X(I)	SUM X(I)
71	141	160	1	1	20	1	162.0	2915.000
72	141	160	1	21	40	1	168.0	3083.000
73	141	160	1	41	60	1	83.0	3166.000
74	141	160	1	61	80	1	40.0	3206.000
75	141	160	1	81	100	1	11.0	3217.000
76	141	160	1	101	120	1	5.0	3222.000
77	141	160	1	121	140	1	196.0	3418.000
78	141	160	1	141	160	1	199.0	3617.000
79	141	160	1	161	180	1	167.0	3784.000
80	141	160	1	181	200	1	195.0	3979.000
81	161	180	1	1	20	1	133.0	4112.000
82	161	180	1	21	40	1	131.0	4243.000
83	161	180	1	41	60	1	203.0	4446.000
84	161	180	1	61	80	1	147.0	4593.000
85	161	180	1	81	100	1	48.0	4641.000
86	161	180	1	101	120	1	31.0	4672.000
87	161	180	1	121	140	1	45.0	4717.000
88	161	180	1	141	160	1	151.0	4868.000
89	161	180	1	161	180	1	242.0	5110.000
90	161	180	1	181	200	1	196.0	5306.000
91	181	200	1	1	20	1	113.0	5419.000
92	181	200	1	21	40	1	79.0	5498.000
93	181	200	1	41	60	1	182.0	5680.000
94	181	200	1	61	80	1	240.0	5920.000
95	181	200	1	81	100	1	217.0	6137.000
96	181	200	1	101	120	1	144.0	6281.000
97	181	200	1	121	140	1	120.0	6401.000
98	181	200	1	141	160	1	24.0	6425.000
99	181	200	1	161	180	1	43.0	6468.000
100	181	200	1	181	200	1	169.0	6637.000

MEAN = 66.37000
STANDARD DEVIATION = 70.81426
COEFFICIENT OF VARIATION = 1.06696

Table 30. Sample list for MEADOW sample using ORIGINAL classification.

PRIMARY SAMPLE UNITS FOR EASTP								
PSU #	LINES			COLS.			X(I)	SUM X(I)
1	1	25	1	1	25	1	54.0	54.000
2	1	25	1	26	50	1	58.0	112.000
3	1	25	1	51	75	1	42.0	154.000
4	1	25	1	76	100	1	275.0	429.000
5	1	25	1	101	125	1	28.0	457.000
6	1	25	1	126	150	1	96.0	553.000
7	1	25	1	151	175	1	56.0	609.000
8	1	25	1	176	200	1	69.0	678.000
9	26	50	1	1	25	1	74.0	752.000
10	26	50	1	26	50	1	85.0	837.000
11	26	50	1	51	75	1	3.0	840.000
12	26	50	1	76	100	1	154.0	994.000
13	26	50	1	101	125	1	77.0	1071.000
14	26	50	1	126	150	1	55.0	1126.000
15	26	50	1	151	175	1	33.0	1159.000
16	26	50	1	176	200	1	66.0	1225.000
17	51	75	1	1	25	1	105.0	1330.000
18	51	75	1	26	50	1	168.0	1498.000
19	51	75	1	51	75	1	56.0	1554.000
20	51	75	1	76	100	1	100.0	1654.000
21	51	75	1	101	125	1	144.0	1798.000
22	51	75	1	126	150	1	28.0	1826.000
23	51	75	1	151	175	1	37.0	1863.000
24	51	75	1	176	200	1	24.0	1887.000
25	76	100	1	1	25	1	130.0	2017.000
26	76	100	1	26	50	1	117.0	2134.000
27	76	100	1	51	75	1	41.0	2175.000
28	76	100	1	76	100	1	24.0	2199.000
29	76	100	1	101	125	1	42.0	2241.000
30	76	100	1	126	150	1	104.0	2345.000
31	76	100	1	151	175	1	89.0	2434.000
32	76	100	1	176	200	1	79.0	2513.000
33	101	125	1	1	25	1	48.0	2561.000
34	101	125	1	26	50	1	97.0	2658.000
35	101	125	1	51	75	1	32.0	2690.000
36	101	125	1	76	100	1	14.0	2704.000
37	101	125	1	101	125	1	16.0	2720.000
38	101	125	1	126	150	1	92.0	2812.000
39	101	125	1	151	175	1	56.0	2868.000
40	101	125	1	176	200	1	25.0	2893.000
41	126	150	1	1	25	1	61.0	2954.000
42	126	150	1	26	50	1	124.0	3078.000
43	126	150	1	51	75	1	66.0	3144.000
44	126	150	1	76	100	1	14.0	3158.000
45	126	150	1	101	125	1	47.0	3205.000
46	126	150	1	126	150	1	14.0	3219.000
47	126	150	1	151	175	1	76.0	3295.000
48	126	150	1	176	200	1	24.0	3319.000
49	151	175	1	1	25	1	78.0	3397.000
50	151	175	1	26	50	1	45.0	3442.000
51	151	175	1	51	75	1	97.0	3539.000
52	151	175	1	76	100	1	122.0	3661.000
53	151	175	1	101	125	1	86.0	3747.000
54	151	175	1	126	150	1	32.0	3779.000
55	151	175	1	151	175	1	18.0	3797.000
56	151	175	1	176	200	1	37.0	3834.000
57	176	200	1	1	25	1	173.0	4007.000
58	176	200	1	26	50	1	160.0	4167.000
59	176	200	1	51	75	1	22.0	4189.000
60	176	200	1	76	100	1	54.0	4243.000
61	176	200	1	101	125	1	66.0	4309.000
62	176	200	1	126	150	1	132.0	4441.000
63	176	200	1	151	175	1	59.0	4500.000
64	176	200	1	176	200	1	26.0	4526.000

Table 30. (Continued)

```
 MEAN =          70.71875
STANDARD DEVIATION =          49.27360
COEFFICIENT OF VARIATION =                .69675
```

Table 31. Sample list for MEADOW sample using JOSESIGS classification.

PRIMARY SAMPLE UNITS FOR PSU =		LINES			COLS.		X(I)	SUM X(I)
1	1	25	1	1	25	1	19.0	19.000
2	1	25	1	26	50	1	29.0	48.000
3	1	25	1	51	75	1	2.0	50.000
4	1	25	1	76	100	1	104.0	154.000
5	1	25	1	101	125	1	0.0	154.000
6	1	25	1	126	150	1	31.0	185.000
7	1	25	1	151	175	1	11.0	196.000
8	1	25	1	176	200	1	38.0	234.000
9	26	50	1	1	25	1	14.0	248.000
10	26	50	1	26	50	1	27.0	275.000
11	26	50	1	51	75	1	1.0	276.000
12	26	50	1	76	100	1	63.0	339.000
13	26	50	1	101	125	1	28.0	367.000
14	26	50	1	126	150	1	25.0	392.000
15	26	50	1	151	175	1	9.0	401.000
16	26	50	1	176	200	1	16.0	417.000
17	51	75	1	1	25	1	61.0	478.000
18	51	75	1	26	50	1	111.0	589.000
19	51	75	1	51	75	1	19.0	608.000
20	51	75	1	76	100	1	60.0	668.000
21	51	75	1	101	125	1	74.0	742.000
22	51	75	1	126	150	1	26.0	768.000
23	51	75	1	151	175	1	9.0	777.000
24	51	75	1	176	200	1	4.0	781.000
25	76	100	1	1	25	1	39.0	820.000
26	76	100	1	26	50	1	29.0	849.000
27	76	100	1	51	75	1	38.0	887.000
28	76	100	1	76	100	1	3.0	890.000
29	76	100	1	101	125	1	3.0	893.000
30	76	100	1	126	150	1	69.0	962.000
31	76	100	1	151	175	1	50.0	1012.000
32	76	100	1	176	200	1	17.0	1029.000
33	101	125	1	1	25	1	57.0	1086.000
34	101	125	1	26	50	1	36.0	1122.000
35	101	125	1	51	75	1	11.0	1133.000
36	101	125	1	76	100	1	0.0	1133.000
37	101	125	1	101	125	1	11.0	1144.000
38	101	125	1	126	150	1	48.0	1192.000
39	101	125	1	151	175	1	25.0	1217.000
40	101	125	1	176	200	1	1.0	1218.000
41	126	150	1	1	25	1	42.0	1260.000
42	126	150	1	26	50	1	58.0	1318.000
43	126	150	1	51	75	1	25.0	1343.000
44	126	150	1	76	100	1	4.0	1347.000
45	126	150	1	101	125	1	24.0	1371.000
46	126	150	1	126	150	1	11.0	1382.000
47	126	150	1	151	175	1	62.0	1444.000
48	126	150	1	176	200	1	22.0	1466.000
49	151	175	1	1	25	1	72.0	1538.000
50	151	175	1	26	50	1	22.0	1560.000
51	151	175	1	51	75	1	104.0	1664.000
52	151	175	1	76	100	1	84.0	1748.000
53	151	175	1	101	125	1	50.0	1798.000
54	151	175	1	126	150	1	3.0	1801.000
55	151	175	1	151	175	1	10.0	1811.000
56	151	175	1	176	200	1	19.0	1830.000
57	176	200	1	1	25	1	101.0	1931.000
58	176	200	1	26	50	1	61.0	1992.000
59	176	200	1	51	75	1	15.0	2007.000
60	176	200	1	76	100	1	27.0	2034.000
61	176	200	1	101	125	1	52.0	2086.000
62	176	200	1	126	150	1	71.0	2157.000
63	176	200	1	151	175	1	45.0	2202.000
64	176	200	1	176	200	1	14.0	2216.000

Table 31. (Continued)

```
 MEAN =          34.62500
STANDARD DEVIATION =          28.71563
COEFFICIENT OF VARIATION =              .82933
```

Table 32. Sample list for MEADOW sample using SCRAMBL1 classification.

PRIMARY SAMPLE UNITS FOR PSU #	LINES			COLS.			Y(I)	SUM X(I)
1	1	25	1	1	25	1	41.0	41.000
2	1	25	1	26	50	1	42.0	83.000
3	1	25	1	51	75	1	34.0	117.000
4	1	25	1	76	100	1	192.0	309.000
5	1	25	1	101	125	1	21.0	330.000
6	1	25	1	126	150	1	83.0	413.000
7	1	25	1	151	175	1	47.0	460.000
8	1	25	1	176	200	1	58.0	518.000
9	26	50	1	1	25	1	57.0	575.000
10	26	50	1	26	50	1	64.0	639.000
11	26	50	1	51	75	1	2.0	641.000
12	26	50	1	76	100	1	116.0	757.000
13	26	50	1	101	125	1	60.0	817.000
14	26	50	1	126	150	1	41.0	858.000
15	26	50	1	151	175	1	22.0	880.000
16	26	50	1	176	200	1	52.0	932.000
17	51	75	1	1	25	1	82.0	1014.000
18	51	75	1	26	50	1	131.0	1145.000
19	51	75	1	51	75	1	41.0	1186.000
20	51	75	1	76	100	1	73.0	1259.000
21	51	75	1	101	125	1	109.0	1368.000
22	51	75	1	126	150	1	20.0	1388.000
23	51	75	1	151	175	1	23.0	1411.000
24	51	75	1	176	200	1	18.0	1429.000
25	76	100	1	1	25	1	98.0	1527.000
26	76	100	1	26	50	1	91.0	1618.000
27	76	100	1	51	75	1	25.0	1643.000
28	76	100	1	76	100	1	19.0	1662.000
29	76	100	1	101	125	1	36.0	1698.000
30	76	100	1	126	150	1	76.0	1774.000
31	76	100	1	151	175	1	60.0	1834.000
32	76	100	1	176	200	1	59.0	1893.000
33	101	125	1	1	25	1	38.0	1931.000
34	101	125	1	26	50	1	71.0	2002.000
35	101	125	1	51	75	1	25.0	2027.000
36	101	125	1	76	100	1	14.0	2041.000
37	101	125	1	101	125	1	15.0	2056.000
38	101	125	1	126	150	1	67.0	2123.000
39	101	125	1	151	175	1	37.0	2160.000
40	101	125	1	176	200	1	23.0	2183.000
41	126	150	1	1	25	1	41.0	2224.000
42	126	150	1	26	50	1	94.0	2318.000
43	126	150	1	51	75	1	48.0	2366.000
44	126	150	1	76	100	1	9.0	2375.000
45	126	150	1	101	125	1	37.0	2412.000
46	126	150	1	126	150	1	17.0	2429.000
47	126	150	1	151	175	1	65.0	2494.000
48	126	150	1	176	200	1	23.0	2517.000
49	151	175	1	1	25	1	65.0	2582.000
50	151	175	1	26	50	1	29.0	2611.000
51	151	175	1	51	75	1	73.0	2684.000
52	151	175	1	76	100	1	86.0	2770.000
53	151	175	1	101	125	1	64.0	2834.000
54	151	175	1	126	150	1	26.0	2860.000
55	151	175	1	151	175	1	15.0	2875.000
56	151	175	1	176	200	1	33.0	2908.000
57	176	200	1	1	25	1	131.0	3039.000
58	176	200	1	26	50	1	119.0	3158.000
59	176	200	1	51	75	1	20.0	3178.000
60	176	200	1	76	100	1	49.0	3227.000
61	176	200	1	101	125	1	48.0	3275.000
62	176	200	1	126	150	1	110.0	3385.000
63	176	200	1	151	175	1	51.0	3436.000
64	176	200	1	176	200	1	16.0	3452.000

Table 32. (Continued)

MEAN =	53.93750
STANDARD DEVIATION =	36.29591
COEFFICIENT OF VARIATION =	.67293

Table 33. Sample list for MEADOW sample using SCRAMBL2 classification.

PRIMARY SAMPLE UNITS FOR PASTR PSU =		LINES			COLS.		X(I)	SUM X(I)
1	1	25	1	1	25	1	39.0	39.000
2	1	25	1	26	50	1	39.0	78.000
3	1	25	1	51	75	1	25.0	103.000
4	1	25	1	76	100	1	173.0	276.000
5	1	25	1	101	125	1	25.0	301.000
6	1	25	1	126	150	1	75.0	376.000
7	1	25	1	151	175	1	49.0	425.000
8	1	25	1	176	200	1	67.0	492.000
9	26	50	1	1	25	1	49.0	541.000
10	26	50	1	26	50	1	59.0	600.000
11	26	50	1	51	75	1	7.0	607.000
12	26	50	1	76	100	1	97.0	704.000
13	26	50	1	101	125	1	53.0	757.000
14	26	50	1	126	150	1	31.0	788.000
15	26	50	1	151	175	1	21.0	809.000
16	26	50	1	176	200	1	40.0	849.000
17	51	75	1	1	25	1	76.0	925.000
18	51	75	1	26	50	1	106.0	1031.000
19	51	75	1	51	75	1	42.0	1073.000
20	51	75	1	76	100	1	63.0	1136.000
21	51	75	1	101	125	1	93.0	1229.000
22	51	75	1	126	150	1	18.0	1247.000
23	51	75	1	151	175	1	23.0	1270.000
24	51	75	1	176	200	1	19.0	1289.000
25	76	100	1	1	25	1	104.0	1393.000
26	76	100	1	26	50	1	82.0	1475.000
27	76	100	1	51	75	1	31.0	1506.000
28	76	100	1	76	100	1	20.0	1526.000
29	76	100	1	101	125	1	34.0	1560.000
30	76	100	1	126	150	1	71.0	1631.000
31	76	100	1	151	175	1	65.0	1696.000
32	76	100	1	176	200	1	61.0	1757.000
33	101	125	1	1	25	1	55.0	1812.000
34	101	125	1	26	50	1	57.0	1869.000
35	101	125	1	51	75	1	25.0	1894.000
36	101	125	1	76	100	1	21.0	1915.000
37	101	125	1	101	125	1	41.0	1956.000
38	101	125	1	126	150	1	61.0	2017.000
39	101	125	1	151	175	1	33.0	2050.000
40	101	125	1	176	200	1	28.0	2078.000
41	126	150	1	1	25	1	54.0	2132.000
42	126	150	1	26	50	1	82.0	2214.000
43	126	150	1	51	75	1	54.0	2268.000
44	126	150	1	76	100	1	11.0	2279.000
45	126	150	1	101	125	1	37.0	2316.000
46	126	150	1	126	150	1	37.0	2353.000
47	126	150	1	151	175	1	79.0	2432.000
48	126	150	1	176	200	1	33.0	2465.000
49	151	175	1	1	25	1	82.0	2547.000
50	151	175	1	26	50	1	39.0	2586.000
51	151	175	1	51	75	1	75.0	2661.000
52	151	175	1	76	100	1	95.0	2756.000
53	151	175	1	101	125	1	56.0	2812.000
54	151	175	1	126	150	1	23.0	2835.000
55	151	175	1	151	175	1	27.0	2862.000
56	151	175	1	176	200	1	50.0	2912.000
57	176	200	1	1	25	1	126.0	3038.000
58	176	200	1	26	50	1	112.0	3150.000
59	176	200	1	51	75	1	25.0	3175.000
60	176	200	1	76	100	1	38.0	3213.000
61	176	200	1	101	125	1	69.0	3282.000
62	176	200	1	126	150	1	101.0	3383.000
63	176	200	1	151	175	1	42.0	3425.000
64	176	200	1	176	200	1	25.0	3450.000

Table 33. (Continued)

```
 MEAN =          53.90625
STANDARD DEVIATION =          31.51000
COEFFICIENT OF VARIATION =              .58453
```

Table 34. Sample list for MEADOW sample using SCRAMBL3 classification.

PRIMARY SAMPLE UNITS FOR PASTO PSU #		LINES			COLS.		Y(I)	SUM Y(I)
1	1	25	1	1	25	1	87.0	87.000
2	1	25	1	26	50	1	70.0	157.000
3	1	25	1	51	75	1	66.0	223.000
4	1	25	1	76	100	1	258.0	481.000
5	1	25	1	101	125	1	53.0	534.000
6	1	25	1	126	150	1	107.0	641.000
7	1	25	1	151	175	1	64.0	705.000
8	1	25	1	176	200	1	76.0	781.000
9	26	50	1	1	25	1	97.0	878.000
10	26	50	1	26	50	1	123.0	1001.000
11	26	50	1	51	75	1	41.0	1042.000
12	26	50	1	76	100	1	163.0	1205.000
13	26	50	1	101	125	1	99.0	1304.000
14	26	50	1	126	150	1	74.0	1378.000
15	26	50	1	151	175	1	55.0	1433.000
16	26	50	1	176	200	1	74.0	1507.000
17	51	75	1	1	25	1	122.0	1629.000
18	51	75	1	26	50	1	171.0	1800.000
19	51	75	1	51	75	1	77.0	1877.000
20	51	75	1	76	100	1	117.0	1994.000
21	51	75	1	101	125	1	148.0	2142.000
22	51	75	1	126	150	1	47.0	2189.000
23	51	75	1	151	175	1	79.0	2268.000
24	51	75	1	176	200	1	55.0	2323.000
25	76	100	1	1	25	1	135.0	2458.000
26	76	100	1	26	50	1	132.0	2590.000
27	76	100	1	51	75	1	60.0	2650.000
28	76	100	1	76	100	1	47.0	2697.000
29	76	100	1	101	125	1	68.0	2765.000
30	76	100	1	126	150	1	112.0	2877.000
31	76	100	1	151	175	1	105.0	2982.000
32	76	100	1	176	200	1	90.0	3072.000
33	101	125	1	1	25	1	62.0	3134.000
34	101	125	1	26	50	1	112.0	3246.000
35	101	125	1	51	75	1	55.0	3301.000
36	101	125	1	76	100	1	36.0	3337.000
37	101	125	1	101	125	1	30.0	3367.000
38	101	125	1	126	150	1	98.0	3465.000
39	101	125	1	151	175	1	73.0	3538.000
40	101	125	1	176	200	1	36.0	3574.000
41	126	150	1	1	25	1	65.0	3639.000
42	126	150	1	26	50	1	137.0	3776.000
43	126	150	1	51	75	1	87.0	3863.000
44	126	150	1	76	100	1	39.0	3902.000
45	126	150	1	101	125	1	64.0	3966.000
46	126	150	1	126	150	1	33.0	3999.000
47	126	150	1	151	175	1	96.0	4095.000
48	126	150	1	176	200	1	39.0	4134.000
49	151	175	1	1	25	1	90.0	4224.000
50	151	175	1	26	50	1	63.0	4287.000
51	151	175	1	51	75	1	107.0	4394.000
52	151	175	1	76	100	1	135.0	4529.000
53	151	175	1	101	125	1	98.0	4627.000
54	151	175	1	126	150	1	43.0	4670.000
55	151	175	1	151	175	1	23.0	4693.000
56	151	175	1	176	200	1	48.0	4741.000
57	176	200	1	1	25	1	167.0	4908.000
58	176	200	1	26	50	1	169.0	5077.000
59	176	200	1	51	75	1	36.0	5113.000
60	176	200	1	76	100	1	62.0	5175.000
61	176	200	1	101	125	1	77.0	5252.000
62	176	200	1	126	150	1	144.0	5396.000
63	176	200	1	151	175	1	73.0	5469.000
64	176	200	1	176	200	1	41.0	5510.000

Table 34. (Continued)

```
 MEAN =          85.03375
STANDARD DEVIATION =          43.95514
COEFFICIENT OF VARIATION =             .51655
```

APPENDIX C

Sample Data

Table 35. Experiment 1 sample data.

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	174	.03014	2.877	95.5
82	113	.01957	2.425	123.9
88	146	.02529	2.123	84.0
89	236	.04087	3.845	94.1
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 99.2 $s^2 = 220.8$ $s = 14.9$

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	225	.02584	2.877	111.3
82	205	.02354	2.425	103.0
88	182	.02090	2.123	101.6
89	288	.03307	3.845	116.3
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 113.1 $s^2 = 164.5$ $s = 12.8$

Table 35. (Continued)

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	165	.02995	2.877	96.1
82	111	.02015	2.425	120.4
88	140	.02541	2.123	83.5
89	223	.04048	3.845	95.0
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 98.6 $s^2 = 179.9$ $s = 13.4$

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	148	.02941	2.877	97.8
82	101	.02007	2.425	120.8
88	120	.02385	2.123	89.0
89	188	.03736	3.845	102.9
94	202	.04014	4.135	103.0

$\hat{A}$ (mi.^{2}) = 102.7 $s^2 = 134.8$ $s = 11.6$

Table 35. (Continued)

SAMPLE DATA FOR: Experiment: 1

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i(mi.^{2})
70	175	.02637	2.877	109.1
82	131	.01974	2.425	122.9
88	151	.02275	2.123	93.3
89	242	.03646	3.845	105.5
94	240	.03616	4.135	114.4

$\hat{A}$ (mi.^{2}) = 109.0 $s^2 = 119.8$ $s = 10.9$

Table 36. Experiment 2 sample data.

SAMPLE DATA FOR: Experiment: 2

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	174	.03014	2.877	95.5
82	113	.01957	2.425	123.9
88	146	.02529	2.123	84.0
89	236	.04087	3.845	94.1
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 99.2 $s^2 = 220.8$ $s = 14.9$

SAMPLE DATA FOR: Experiment: 2

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
67	211	.02423	3.152	130.1
80	253	.02905	3.189	109.8
84	183	.02102	1.912	91.0
89	288	.03307	3.845	116.3
93	246	.02825	3.394	120.1

$\hat{A}$ (mi.^{2}) = 113.4 $s^2 = 211.9$ $s = 14.6$

Table 36. (Continued)

SAMPLE DATA FOR: Experiment: 2

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	165	.02995	2.877	96.1
82	111	.02015	2.425	120.4
88	140	.02541	2.123	83.5
89	223	.04048	3.845	95.0
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 98.6 $s^2 = 179.9$ $s = 13.4$

SAMPLE DATA FOR: Experiment: 2

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
69	76	.01510	1.694	112.2
81	97	.01928	2.302	119.4
88	120	.02385	2.123	89.0
89	188	.03736	3.845	102.9
93	143	.02842	3.394	119.4

$\hat{A}$ (mi.^{2}) = 108.6 $s^2 = 165.6$ $s = 12.9$

Table 36. (Continued)

SAMPLE DATA FOR: Experiment: 2

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
67	115	.01733	3.152	181.9
80	195	.02938	3.189	108.5
86	31	.00467	0.127	27.2
89	242	.03646	3.845	105.5
93	182	.02742	3.394	123.8

$\hat{A}$ (mi.^{2}) = 109.4 $s^2 = 3059.8$ $s = 55.3$

Table 37. Experiment 3 sample data.

SAMPLE DATA FOR: Experiment: 3

Land Cover Class: MEADOW Classification: ORIGINAL

$n = 9$ $N = 64$ $\sum_{i=1}^{N} x_i = 4526$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
4	275	.06076	2.948	48.5
14	55	.01215	0.618	50.9
15	33	.00729	0.084	11.5
18	168	.03712	2.168	58.4
25	130	.02872	0.340	11.8
43	66	.01458	0.900	61.7
49	78	.01723	0.357	20.7
57	173	.03822	1.117	29.2
63	59	.01304	0.704	54.0

$\hat{A}$ (mi.^{2}) = 38.5 $s^2 = 408.7$ $s = 20.2$

Table 37. (Continued)

SAMPLE DATA FOR: Experiment: 3

Land Cover Class: MEADOW Classification: JOSESIGS

$n = 9$ $N = 64$ $\sum_{i=1}^{N} x_i = 2216$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
4	104	.04693	2.948	62.8
14	25	.01128	0.618	54.8
15	9	.00406	0.084	20.7
18	111	.05009	2.168	43.3
25	39	.01760	0.340	19.3
43	25	.01128	0.900	79.8
49	72	.03249	0.357	11.0
57	101	.04558	1.117	24.5
63	45	.02031	0.704	34.7

$\hat{A}$ (mi.^{2}) = 39.0 $s^2 = 529.1$ $s = 23.0$

Table 37. (Continued)

SAMPLE DATA FOR: Experiment: 3

Land Cover Class: MEADOW Classification: SCRAMBL1

$n = 9$ $N = 64$ $\sum_{i=1}^{N} x_i = 3452$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
4	192	.05562	2.948	53.0
14	41	.01188	0.618	52.0
15	22	.00637	0.084	13.2
18	131	.03795	2.168	57.1
25	98	.02839	0.340	12.0
43	48	.01390	0.900	64.7
49	65	.01883	0.357	19.0
57	131	.03795	1.117	29.4
63	51	.01477	0.704	47.7

$\hat{A}$ (mi.^{2}) = 38.7 $s^2 = 415.0$ $s = 20.4$

Table 37. (Continued)

SAMPLE DATA FOR: Experiment: 3

Land Cover Class: MEADOW Classification: SCRAMBL2

$n = 9$ $N = 64$ $\sum_{i=1}^{N} x_i = 3450$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
4	173	.05014	2.948	58.8
14	31	.00899	0.618	68.8
15	21	.00609	0.084	13.8
18	106	.03072	2.168	70.6
25	104	.03014	0.340	11.3
43	54	.01565	0.900	57.5
49	82	.02377	0.357	15.0
57	126	.03652	1.117	30.6
63	42	.01217	0.704	57.8

$\hat{A}$ (mi.^{2}) = 42.7 $s^2 = 612.3$ $s = 24.7$

Table 37. (Continued)

SAMPLE DATA FOR: Experiment: 3

Land Cover Class: MEADOW Classification: SCRAMBL3

$n = 9$ $N = 64$ $\sum_{i=1}^{N} x_i = 5510$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
4	258	.04682	2.948	63.0
14	74	.01343	0.618	46.0
15	55	.00998	0.084	8.4
18	171	.03103	2.168	69.9
25	135	.02450	0.340	13.9
43	87	.01579	0.900	57.0
49	90	.01633	0.357	21.9
57	167	.03031	1.117	36.9
63	73	.01325	0.704	53.1

$\hat{A}$ (mi.^{2}) = 41.1 $s^2 = 490.5$ $s = 22.1$

Table 38. Experiment 4 sample data.

SAMPLE DATA FOR: Experiment: 4 - 1

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	174	.03014	2.877	95.5
82	113	.01957	2.425	123.9
88	146	.02529	2.123	84.0
89	236	.04087	3.845	94.1
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 99.2 $s^2 = 220.8$ $s = 14.9$

SAMPLE DATA FOR: Experiment: 4 - 1

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	225	.02584	2.877	111.3
82	205	.02354	2.425	103.0
88	182	.02090	2.123	101.6
89	288	.03307	3.845	116.3
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 113.1 $s^2 = 164.5$ $s = 12.8$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 1

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	165	.02995	2.877	96.1
82	111	.02015	2.425	120.4
88	140	.02541	2.123	83.5
89	223	.04048	3.845	95.0
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 98.6 $s^2 = 179.9$ $s = 13.4$

SAMPLE DATA FOR: Experiment: 4 - 1

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	148	.02941	2.877	97.8
82	101	.02007	2.425	120.8
88	120	.02385	2.123	89.0
89	188	.03736	3.845	102.9
94	202	.04014	4.135	103.0

$\hat{A}$ (mi.^{2}) = 102.7 $s^2 = 134.8$ $s = 11.6$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 1

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
70	175	.02637	2.877	109.1
82	131	.01974	2.425	122.9
88	151	.02275	2.123	93.3
89	242	.03646	3.845	105.5
94	240	.03616	4.135	114.4

$\hat{A}$ (mi.^{2}) = 109.0 $s^2 = 119.8$ $s = 10.9$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 2

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	106	.01836	1.431	77.9
59	40	.00693	0.599	86.5
70	174	.03014	2.877	95.5
83	190	.03291	3.087	93.8
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 90.5 $s^2 = 69.0$ $s = 8.3$

SAMPLE DATA FOR: Experiment: 4 - 2

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	118	.01355	1.431	105.6
59	48	.00551	0.599	108.7
70	225	.02584	2.877	111.3
83	240	.02756	3.087	112.0
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 114.2 $s^2 = 121.1$ $s = 11.0$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 2

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	102	.01852	1.431	77.3
59	38	.00690	0.599	86.8
70	165	.02995	2.877	96.1
83	178	.03231	3.087	95.5
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 90.8 $s^2 = 75.7$ $s = 8.7$

SAMPLE DATA FOR: Experiment: 4 - 2

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	88	.01749	1.431	81.8
59	37	.00735	0.599	81.5
70	148	.02941	2.877	97.8
83	154	.03060	3.087	100.9
94	202	.04014	4.135	103.0

$\hat{A}$ (mi.^{2}) = 93.0 $s^2 = 110.8$ $s = 10.5$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 2

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	105	.01582	1.431	90.5
59	47	.00708	0.599	84.6
70	175	.02637	2.877	109.1
83	203	.03059	3.087	100.9
94	240	.03616	4.135	114.4

$\hat{A}$ (mi.^{2}) = 99.9 $s^2 = 154.6$ $s = 12.4$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 3

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	106	.01836	1.431	77.9
51	95	.01645	3.236	196.7
67	91	.01576	3.152	200.0
69	85	.01472	1.694	115.1
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 137.7 s^2 = 3241.6 s = 56.9

SAMPLE DATA FOR: Experiment: 4 - 3

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	118	.01355	1.431	105.6
51	195	.02239	3.236	144.5
67	211	.02423	3.152	130.1
69	118	.01355	1.694	125.0
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 127.7 s^2 = 204.0 s = 14.3

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 3

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i(mi.^{2})
19	102	.01852	1.431	77.3
51	90	.01634	3.236	198.1
67	88	.01597	3.152	197.3
69	82	.01488	1.694	113.8
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 136.9 $s^2 = 3244.6$ $s = 56.0$

SAMPLE DATA FOR: Experiment: 4 - 3

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i(mi.^{2})
19	88	.01749	1.431	81.8
51	88	.01749	3.236	185.0
67	88	.01749	3.152	180.2
69	76	.01510	1.694	112.2
94	202	.04014	4.135	103.0

$\hat{A}$ (mi.^{2}) = 132.5 $s^2 = 2222.6$ $s = 47.1$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 3

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	105	.01582	1.431	90.5
51	121	.01823	3.236	177.5
67	115	.01733	3.152	181.9
69	98	.01477	1.694	114.7
94	240	.03616	4.135	114.4

$\hat{A}$ (mi.^{2}) = 135.8 $s^2 = 1706.3$ $s = 41.3$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 4

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	P_i	y_i	A_i(mi.^{2})
59	40	.00693	0.599	86.5
70	174	.03014	2.877	95.5
80	172	.02979	3.189	107.1
81	121	.02096	2.302	109.8
83	190	.03291	3.087	93.8

$\hat{A}$ (mi.^{2}) = 98.5 $s^2 = 94.5$ $s = 9.7$

SAMPLE DATA FOR: Experiment: 4 - 4

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	P_i	y_i	A_i(mi.^{2})
59	48	.00551	0.599	108.7
70	225	.02584	2.877	111.3
80	253	.02905	3.189	109.8
81	178	.02044	2.302	112.6
83	240	.02756	3.087	112.0

$\hat{A}$ (mi.^{2}) = 110.9 $s^2 = 2.7$ $s = 1.6$

Table 38. (Continued)
SAMPLE DATA FOR: Experiment: 4 - 4

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
59	38	.00690	0.599	86.8
70	165	.02995	2.877	96.1
80	167	.03031	3.189	105.2
81	114	.02069	2.302	111.2
83	178	.03231	3.087	95.5

$\hat{A}$ (mi.^{2}) = 99.0 $s^2 = 89.2$ $s = 9.4$

SAMPLE DATA FOR: Experiment: 4 - 4

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
59	37	.00735	0.599	81.5
70	148	.02941	2.877	97.8
80	139	.02762	3.189	115.4
81	97	.01928	2.302	119.4
83	154	.03060	3.087	100.9

$\hat{A}$ (mi.^{2}) = 103.0 $s^2 = 229.1$ $s = 15.1$

Table 38. (Continued).

SAMPLE DATA FOR: Experiment: 4 - 4

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
59	47	.00708	0.599	84.6
70	175	.02637	2.877	109.1
80	195	.02938	3.189	108.5
81	133	.02004	2.302	114.9
83	203	.03059	3.087	100.9

$\hat{A}$ (mi.^{2}) = 103.6 $s^2 = 138.1$ $s = 11.8$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 5

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
40	43	.00745	0.615	82.6
60	153	.02650	2.714	102.4
62	93	.01611	2.573	159.7
69	85	.01472	1.694	115.1
80	172	.02979	3.189	107.1

$\hat{A}$ (mi.^{2}) = 113.4 $s^2 = 814.0$ $s = 28.5$

SAMPLE DATA FOR: Experiment: 4 - 5

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
40	55	.00632	0.615	97.4
60	230	.02641	2.714	102.8
62	158	.01814	2.573	141.8
69	118	.01355	1.694	125.0
80	253	.02905	3.189	109.8

$\hat{A}$ (mi.^{2}) = 115.4 $s^2 = 325.8$ $s = 18.1$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 5

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
40	41	.00744	0.615	82.6
60	149	.02705	2.714	100.3
62	90	.01634	2.573	157.5
69	82	.01488	1.694	113.8
80	167	.03031	3.189	105.2

$\hat{A}$ (mi.^{2}) = 111.9 $s^2 = 780.2$ $s = 27.9$

SAMPLE DATA FOR: Experiment: 4 - 5

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
40	36	.00715	0.615	86.0
60	139	.02762	2.714	98.3
62	84	.01669	2.573	154.1
69	76	.01510	1.694	112.2
80	139	.02762	3.189	115.4

$\hat{A}$ (mi.^{2}) = 113.2 $s^2 = 660.1$ $s = 25.7$

Table 38. (Continued)

SAMPLE DATA FOR: Experiment: 4 - 5

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 5$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
40	44	.00663	0.615	92.8
60	170	.02561	2.714	106.0
62	106	.01597	2.573	161.1
69	98	.01477	1.694	114.7
80	195	.02938	3.189	108.5

$\hat{A}$ (mi.^{2}) = 116.6 $s^2 = 682.1$ $s = 26.1$

Table 39. Experiment 5 sample data.

SAMPLE DATA FOR: Experiment: 5

Land Cover Class: CONIFER Classification: ORIGINAL

$n = 19$ $N = 100$ $\sum_{i=1}^{N} x_i = 5774$

Sample Unit	x_i	p_i	y_i	A_i(mi.^{2})
19	106	.01836	1.431	78.0
40	43	.00745	0.615	82.6
51	95	.01645	3.236	196.7
59	40	.00693	0.599	86.5
60	153	.02650	2.714	102.4
62	93	.01611	2.573	159.7
67	91	.01576	3.152	200.0
69	85	.01472	1.694	115.1
70	174	.03014	2.877	95.5
80	172	.02979	3.189	107.1
81	121	.02096	2.302	109.8
82	113	.01957	2.425	123.9
83	190	.03291	3.087	93.8
84	134	.02321	1.912	82.4
86	22	.00381	0.127	33.3
88	146	.02529	2.123	84.0
89	236	.04087	3.845	94.1
93	177	.03065	3.394	110.7
94	242	.04191	4.135	98.7

$\hat{A}$ (mi.^{2}) = 108.1 s^2 = 1600.4 s = 40.0

Table 39. (Continued)
SAMPLE DATA FOR: Experiment: 5

Land Cover Class: CONIFER Classification: JOSESIGS

$n = 19$ $N = 100$ $\sum_{i=1}^{N} x_i = 8708$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	118	.01355	1.431	105.6
40	55	.00632	0.615	97.4
51	195	.02239	3.236	144.5
59	48	.00551	0.599	108.7
60	230	.02641	2.714	102.8
62	158	.01814	2.573	141.8
67	211	.02423	3.152	130.1
69	118	.01355	1.694	125.0
70	225	.02584	2.877	111.3
80	253	.02905	3.189	109.8
81	178	.02044	2.302	112.6
82	205	.02354	2.425	103.0
83	240	.02756	3.087	112.0
84	183	.02102	1.912	91.0
86	43	.00494	0.127	25.7
88	182	.02090	2.123	101.6
89	288	.03307	3.845	116.3
93	246	.02825	3.394	120.1
94	270	.03101	4.135	133.4

$\hat{A}$ (mi.^{2}) = 110.1 $s^2 = 630.6$ $s = 25.1$

Table 39. (Continued)

SAMPLE DATA FOR: Experiment: 5

Land Cover Class: CONIFER Classification: SCRAMBL1

$n = 19$ $N = 100$ $\sum_{i=1}^{N} x_i = 5509$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	102	.01852	1.431	77.3
40	41	.00744	0.615	82.6
51	90	.01634	3.236	198.1
59	38	.00690	0.599	86.8
60	149	.02705	2.714	100.3
62	90	.01634	2.573	157.5
67	88	.01597	3.152	197.3
69	82	.01488	1.694	113.8
70	165	.02995	2.877	96.1
80	167	.03031	3.189	105.2
81	114	.02069	2.302	111.2
82	111	.02015	2.425	120.4
83	178	.03231	3.087	95.6
84	127	.02305	1.912	82.9
86	22	.00399	0.127	31.8
88	140	.02541	2.123	83.5
89	223	.04048	3.845	95.0
93	168	.03050	3.394	111.3
94	232	.04211	4.135	98.2

$\hat{A}$ (mi.^{2}) = 107.6 $s^2 = 1580.6$ $s = 39.8$

Table 39. (Continued)
SAMPLE DATA FOR: Experiment: 5

Land Cover Class: CONIFER Classification: SCRAMBL2

$n = 19$ $N = 100$ $\sum_{i=1}^{N} x_i = 5032$

Sample Unit	x_i	p_i	y_i	A_i (mi.^{2})
19	88	.01749	1.431	81.8
40	36	.00715	0.615	86.0
51	88	.01749	3.236	185.0
59	37	.00735	0.599	81.5
60	139	.02762	2.714	98.3
62	84	.01669	2.573	154.1
67	88	.01749	3.152	180.2
69	76	.01510	1.694	112.2
70	148	.02941	2.877	97.8
80	139	.02762	3.189	115.4
81	97	.01928	2.302	119.4
82	101	.02007	2.425	120.8
83	154	.03060	3.087	100.9
84	117	.02325	1.912	82.2
86	25	.00497	0.127	25.6
88	120	.02385	2.123	89.0
89	188	.03736	3.845	102.9
93	143	.02842	3.394	119.4
94	202	.04104	4.135	103.0

$\hat{A}$ (mi.^{2}) = 108.2 s^2 = 1319.5 s = 36.3

Table 39. (Continued)

SAMPLE DATA FOR: Experiment: 5

Land Cover Class: CONIFER Classification: SCRAMBL3

$n = 19$ $N = 100$ $\sum_{i=1}^{N} x_i = 6637$

Sample Unit	x_i	p_i	y_i	A_i(mi.^{2})
19	105	.01582	1.431	90.5
40	44	.00663	0.615	92.8
51	121	.01823	3.236	177.5
59	47	.00708	0.599	84.6
60	170	.02561	2.714	106.0
62	106	.01597	2.573	161.1
67	115	.01733	3.152	181.9
69	98	.01477	1.694	114.7
70	175	.02637	2.877	109.1
80	195	.02938	3.189	108.5
81	133	.02004	2.302	114.9
82	131	.01974	2.425	122.9
83	203	.03059	3.087	100.9
84	147	.02215	1.912	86.3
86	31	.00467	0.127	27.2
88	151	.02275	2.123	93.3
89	242	.03646	3.845	105.5
93	182	.02742	3.394	123.8
94	240	.03616	4.135	114.3

$\hat{A}$ (mi.^{2}) = 111.4 s^2 = 1214.6 s = 34.9

Appendix D

PSU and SCRAMBL
Computer Program Lists

```
PROGRAM PSU        73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

  1                 PROGRAM PSU(INPUT,OUTPUT,TAPE4=OUTPUT,TAPE5=INPUT,TAPE7,TAPE8)
                    COMMON /INPUT/ ICARD(80),SCREEN
                    COMMON /BUF/ IA(500)
                    COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
  5                +TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                C
                C
                C
                    INTEGER TABLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,BIN,GETNUM
 10                 C
                C
                C
                    CALL HEADER
                    CALL INFO (4HLINE,SLINE,ELINE,BYLINE,4,3,5,NLINE)
 15                 CALL INFO (4HCOL ,SCOL,ECOL,BYCOL,7,6,8,NCOL)
                    CALL PSU_SIZE
                    CALL CLASIF
                    CALL QUIZ
                    CALL CORE
 20                 END
```

```
SUBROUTINE CORE       72/73   OPT=1                    FTN 4.6+439        79/09/21. 14.34.57

      1               SUBROUTINE CORE
                      COMMON /STATS/ SUM,SQM,IPSU
                      COMMON /BUF/ A (500)
                      COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
      5              +TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                      EQUIVALENCE (A(1),COL(1))
                      INTEGER COL(500),SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,TABLE
                C
                C
     10               ICOUNT = 0
                      SUM = SQM = 0.0
                      IPSU = 0
                      K = SLINE
                      ILINE = 0
     15               LINE = 100
                C
                C
                      DO 200 I = 1,NBINS
                          BIN(I) = 0.0
     20          200  CONTINUE
                C
                C
                      DO 300 I = SLINE , ELINE , BYLINE
                C
     25         C
                        DO 301 J = 1 , K
                           CALL GET LINE
                 301    CONTINUE
                        K = BYLINE
     30                 DO 302 L = SCOL , ECOL , BYCOL
                C
                C
                             DO 303 M = 1 , NCLASS
                                IF ( COL(L).EQ.TABLE(M) ) GO TO 304
     35          303        CONTINUE
                                GO TO 302
                 304          IBIN = NBINS - ( ECOL - L ) / (  IWIDTH )
                              BIN(IBIN) = BIN(IBIN) + 1.0
                 302    CONTINUE
     40         C
                C
                       ILINE = ILINE + BYLINE
                       IF ( ILINE.NE.ILENGHT ) GO TO 300
                          ILINE = 0
     45                   ICOUNT = ICOUNT + 1
                          CALL OUTPUT ( ICOUNT,LINE )
                C
                C
                             DO 310 II = 1,NBINS
     50                         BIN(II) = 0.0
                 310         CONTINUE
                C
                 300  CONTINUE
                C
     55         C
                      CALL STATS
                      RETURN
                      END
```

```
SUBROUTINE INFO       73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

 1                SUBROUTINE INFO (TYPE,S,E,BY,L,M,N,NU)
                  COMMON /INPUT/ ICARD(80),SCREEN
                  COMMON /BUF/ IA(500)
                  INTEGER S,E,BY,GETNUM
 5          C
            C
            C
             100  CONTINUE
                  WRITE(6,600) TYPE
10           600  FORMAT(* PLEASE ENTER *,A4,* INFORMATION CONCERNING THE *
                 +*MAP-FILE  */)
                  READ(5,500) ICARD
             500  FORMAT(80R1)
                  IF (EOF(5).NE.0) GO TO 700
15                I = 1
                  S = GET NUM (I)
                  E = GET NUM (I)
                  BY = GET NUM (I)
                  IF (S.EQ.0.OR.E.EQ.0.OR.BY.EQ.0) GO TO 200
20                WRITE(6,605) S,E,BY
             605   FORMAT(1X,I7,2(3H , ,I7))
                     NU = FLOAT(E - S + 1)
                     NF = FLOAT(IA(L) - IA(M) + 1)/FLOAT(IA(N)) + .99
                       IF (NF.LT.NU) GO TO 250
25                       RETURN
            C
            C
             200  WRITE(6,610) TYPE
             610  FORMAT(* INCORRECT *,A4,* INPUT PARAMETER -RE ENTER-*/)
30                  GO TO 100
            C
            C
             250  WRITE(6,620) NF,NU
             620  FORMAT(* INPUT DIMENSIONS EXCEED MAP-FILE, MAP-FILE CONTAINS*/
                 +1X,I6,* POINTS, YOUR ENTRY WAS *,I7,* POINTS*/)
35                  GO TO 100
            C
            C
             700  WRITE(6,630)
             630  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT */)
                  RETURN
                  END
```

```
SUBROUTINE PSUSIZE    73/73   OPT=1                    FTN 4.6+439        73/09/21. 14.34.57

 1                SUBROUTINE PSU SIZE
                  COMMON /INPUT/ ICARD(90),SCREEN
                  COMMON /BUF/ IA(500)
                  COMMON /DIM/ SLINE,ELINE,RYLINE,SCOL,ECOL,RYCOL,IPSUL,IPSUC,
 5               *TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                  INTEGER TABLE,SLINE,ELINE,RYLINE,SCOL,ECOL,RYCOL,BIN,GETNUM
                C
                C
                C
10               100  WRITE(6,600)
                 600  FORMAT(* PLEASE ENTER THE PSU DIMENSIONS */)
                      READ(5,500) ICARD
                 500  FORMAT(80R1)
                      IF (EOF(5).NE.0) GO TO 700
15                        I = 1
                          IPSUL = GET NUM (I)
                          IPSUC = GET NUM (I)
                          IF(IPSUL.EQ.0.OR.IPSUC.EQ.0) GO TO 200
                          WRITE(6,605) IPSUL,IPSUC
20               605           FORMAT(1X,I7,3H , I7)
                               IF(MOD(NLINE,IPSUL).NE.0) GO TO 250
                               IF(MOD(NCOL,IPSUC).NE.0) GO TO 250
                               NBINS = NCOL/IPSUC
                                    RETURN
25               C
                 C
                  200  WRITE(6,610)
                  610  FORMAT(* INCORRECT DIMENSIONS FOR PSU -RE ENTER-*/)
                           GO TO 100
30               C
                 C
                  250  WRITE(6,620)
                 620   FORMAT(* PSU DIMENSIONS AS ENTERED ARE NOT AN INTEGRAL *,/,1X,
                      +*NUMBER OF THE FILE DIMENSIONS PREVICUSLY INPUT*/)
35                         GO TO 100
                  700  WRITE(6,607)
                  607  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT *)
                        STOP 6000
                       END
```

```
SUBROUTINE CLASIF     73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

 1                SUBROUTINE CLASIF
                  COMMON /INPUT/ ICARD(80),SCREEN
                  COMMON /BUF/ IA(500)
                  COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
 5               *TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                  INTEGER TABLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,BIN,GETNUM
                C
                C
                 100  WRITE(6,600)
10               600  FORMAT(* PLEASE ENTER CLASSIFICATION INFORMATION*/)
                      READ(5,500) ICARD
                      FORMAT(80R1)
                      IF (EOF(5).NE.0) GO TO 700
                          NCLASS = 0
15                        I = 1
                          DO 200 J = 1,10,1
                              NEXT = NAME(I)
                                   DO 210 K = 1, NCLS, 2
                                     IF(NEXT.EQ.IA(K+8)) GO TO 220
20                  210              CONTINUE
                                          IF (I.GE.80) GO TO 250
                                             GO TO 300
                      220  CONTINUE
                           NAMES(J) = NEXT
25                         TABLE(J) = IA(K+9)
                           NCLASS = NCLASS + 1
                      200  CONTINUE
                C
                      250  RETURN
30              C
                C
                      300  WRITE(6,640) NEXT,(IA(L+8),L=1,NCLS)
                      640  FORMAT(* INCORRECT CLASS NAME *,A10,*THIS MAP-FILE CONTAINS*
                          +/,(5(1X,A10,3X,A1)))
35                    GO TO 100
                      700  WRITE(6,650)
                      650  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT *)
                           RETURN
                           END
```

```
SUBROUTINE QUIZ       73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

 1                    SUBROUTINE QUIZ
                      COMMON /INPUT/ CARD(80),SCREEN
                      LOGICAL SCREEN
                C
 5              C
                      SCREEN = .F.
                      WRITE(6,600)
                600   FORMAT(1X,'WOULD YOU LIKE A DETAILED LISTING OF EACH PSU ')
                      READ(5,500) CARD
10                     IF (EOF(5).NE.0) GO TO 770
                600   FORMAT(80A1)
                      I = 1
                       IF ( NAME(I).EQ.3HYES) SCREEN = .TRUE.
                770   RETURN
15                    END
```

```
SUBROUTINE OUTPUT     73/73   OPT=1                    FTN 4.6+439        75/09/21. 14.34.57

      1               SUBROUTINE OUTPUT ( ICOUNT,LINE )
                      COMMON /INPUT/ CARD(80),SCREEN
                      COMMON /STATS/ SUM,SQM,IPSU
                      COMMON /DIM/ SLINE,FLINE,BYLINE,SCOL,FCOL,BYCOL,ILENGHT,IWIDTH,
      5              +TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                      INTEGER BYCOL,BYLINE,SCOL,SLINE
                      LOGICAL SCREEN
                C
                C
     10         C
                      LINE = LINE + NBINS
                      IF ( LINE.LE.60 ) GO TO 200
                         IF ( SCREEN ) WRITE(6,599) (NAMES(J),J=1,NCLASS)
                         IF ( SCREEN ) WRITE(6,600)
     15         C
                C
                      WRITE(6,599)  (NAMES(J),J=1,NCLASS)
                599   FORMAT(*1*,5X,*PRIMARY SAMPLE UNITS FOR   *,5A10,/,33X,
                     +5A10,/)
     20               WRITE(6,600)
                600   FORMAT (5X,*PSU #*,11X,*LINES*,17X,*COLS.*,11X,*X(I) *,
                     +5X,* SUM  X(I)*,/)
                      LINE = 0
                C
     25         C
                C
                200   IM =  ILENGHT
                      IN =  IWIDTH
                      IFL = SLINE + ( ICOUNT * IM ) - IM
     30               ILL = IM - BYLINE  + IFL
                      IFC = SCOL - IN
                C
                C
                C
     35               DO 300 I = 1 , NBINS
                         SUM = SUM + BIN(I)
                         SQM = SQM + BIN(I)**2
                         IPSU = IPSU + 1
                         IFC = IFC + IM
     40                  ILC = IFC + IN - BYCOL
                             IF ( SCREEN ) WRITE(6,660)IPSU,IFL,ILL,BYLINE,IFC,ILC
                     +,BYCOL,BIN(I),SUM
                            WRITE(3,660) IPSU,IFL,ILL,BYLINE,IFC,ILC,BYCOL,BIN(I),SUM
                660   FORMAT(5X,I5,5X,I5,1X,I5,1X,I5,5X,I5,1X,I5,1X,I5,5X,F5.1,5X,F10.3)
     45             300   CONTINUE
                      RETURN
                      END
```

```
SUBROUTINE GETLINE    73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

 1                SUBROUTINE GET LINE
                  COMMON /BUF/ A (500)
                  COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
                 +TABLE(10),NAMES(10),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
 5                EQUIVALENCE ( A(1),COL(1) )
                  INTEGER COL(250),A
           C
           C
           C
10                BUFFER IN (7,1)  (A(1),A(500))
                  IF (UNIT(7)) 111,770,770
           111    N = LENGTH(7) / 2
                  DECODE ( 5000,10,A(1) ) (COL(L),L=1,N)
           10     FORMAT(250(A1,19X))
15                RETURN
           C
           C
           C
                  ENTRY HEADER
20         C
           C
                  BUFFER IN (7,1) ( A(1),A(500) )
                  IF (UNIT(7)) 110,770,770
           110    N = LENGTH(7)
25                INCR = 15
                  IF (N.GE.100) INCR = 50
                  IK= 8 + INCR * 2
                  NCLS = 0
                  DO 200 J = 9,IK
30                   IF (A(J).NE.0) NCLS = NCLS + 1
           200    CONTINUE
                  WRITE(6,600)(A(J),J=1,8),(A(K+8),K=1,NCLS)
                  WRITE(6,600)(A(J),J=1,8),(A(K+8),K=1,NCLS)
           600    FORMAT(*1   HEADER INFORMATION FROM MAP FILE  *,//,
35                +* NUMBER OF CHANNELS  *, I5,/,* NUMBER OF COLUMNS  *,I5,
                  +/,* LINE INFORMATION  *,I7,2(1H,,I7)/,
                  +* COLUMN INFORMATION  *,I7,2(1H,,I7),/,
                  +* CLASS NAMES AND SYMBOLS  *,/,(9(3X,A10,1X,A1)))
                    N = A(1)
40                WRITE(9,610) (A(J+IK),J=1,N,1)
                  WRITE(6,610) (A(J+IK),J=1,N,1)
           610    FORMAT(/* CHANNELS REPRESENTED IN MAP-FILE */,(8(1X,I4))/)
                  RETURN
           770    WRITE(6,701)
45         701    FORMAT(* END-OF-FILE ENCOUNTERED ON THE MAP-FILE *)
                  STOP 7000
                  END
```

```
FUNCTION NAME        73/73   OPT=1                    FTN 4.6+439        79/09/21. 14.30.51

 1                  INTEGER FUNCTION NAME (ICOL)
                    COMMON /INPUT/ ICARD(80),SCREEN
                C
                C
 5              C
                    NAME = 10H
                    ICHAR = 0
                    DO 100 I = ICOL , 80 , 1
                C
10                      IF(ICARD(I).EQ.55B) GO TO 100
                        IF(ICARD(I).GT.0.AND.ICARD(I).LE.44B) GO TO 200
                        IF(ICARD(I).GT.44B) GO TO 300
                 100    CONTINUE
                C
15                  I = 79
                  300   ICOL = I + 1
                    J = (10 - ICHAR) * 6
                    NAME = SHIFT(NAME,J)
                    RETURN
20              C
                C
                C
                  200   ICHAR = ICHAR + 1
                    NAME = SHIFT (NAME,6).AND.77777777777777777700B
25                  NAME = NAME .OR. ICARD(I)
                    GO TO 100
                    END
```

```
FUNCTION GETNUM     73/73   OPT=1                    FTN 4.6+439        79/09/21. 14.30.51

 1                INTEGER FUNCTION GET NUM (ICOL)
                  COMMON /INPUT/ ICARD(80),SCREEN
                C
                C
 5                GET NUM = 0
                C
                C
                  DO 100 I = ICOL , 80 , 1
                      IF (ICARD(I).EQ.558) GO TO 100
10                    IF (ICARD(I).GE.338.AND.ICARD(I).LE.448) GO TO 200
                      IF (ICARD(I).GT.448) GO TO 300
                 100  CONTINUE
                      ICOL = 80
                      RETURN
15              C
                C
                 200  GET NUM = GET NUM * 10 + ICARD(I) - 338
                      GO TO 100
                 300  ICOL = I + 1
20                    RETURN
                          END
```

```
SUBROUTINE STATS      73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.34.57

 1                    SUBROUTINE STATS
                      COMMON /STATS/ SUM,SQM,IPSU
                C
                C
 5                    XMEAN = SUM / FLOAT(IPSU)
                      IF (IPSU.LE.1) STOP 7777
                      STD = SQRT (( SQM-(SUM**2/FLOAT(IPSU))) / FLOAT (IPSU-1) )
                      CV = STD / XMEAN
                      WRITE(6,600) XMEAN,STD,CV
10                    WRITE(8,600) XMEAN,STD,CV
                600   FORMAT(*1*,5X,*MEAN = *,F15.5,/5X,*STANDARD DEVIATION = *,
                     +F15.5,/5X,*COEFFICIENT OF VARIATION = *,F15.5)
                      RETURN
                      END
```

```
PROGRAM SCRMBL      73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.33.23

      1               PROGRAM SCRMBL(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT,TAPE7,TAPE8,
                     + TAPE9)
                C
                C  PROGRAM TO RECLASSIFY A FILE OF CLASSIFIED DATA BASED ON ESTIMATES OF
      5         C  CLASSIFICATION ACCURACY OF TRAINING OR TEST SETS
                C
                     COMMON /INPUT/ ICARD(80)
                     COMMON /BUFR/ A(1000),LENBUF
                     COMMON /DIMS/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,
     10              *TABLE(20),NCLASS,NCLS,NLINE,NCOL
                     COMMON /CMAT/ CONFUS(20,20)
                C
                     EQUIVALENCE (A(1),COL(1))
                     INTEGER TABLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,GETNUM,COL(1000)
     15         C
                     WRITE(6,1)
                   1 FORMAT(* ENTER RUN NUMBER (FORMAT(I2))*)
                     READ(5,2)N
                   2 FORMAT(I2)
     20              N = 2 * N - 1
                     CALL RANSET(N)
                     CALL HEADER
                     CALL INFO (4HLINE,SLINE,ELINE,BYLINE,4,3,5,NLINE)
                     CALL INFO (4HCOL ,SCOL,ECOL,BYCOL,7,6,8,NCOL)
     25              CALL CLASIF
                     CALL ACCUMC
                     CALL CORE
                     STOP
                     END
```

```
SUBROUTINE CORE       73/73   OPT=1                    FTN 4.6+439       79/09/21. 14.33.23

 1                SUBROUTINE CORE
                  COMMON /BUFR/ COL(1000),LENBUF
                  COMMON /DIMS/ SLINE,ELINE,RYLINE,SCOL,ECOL,RYCOL,TABLE(20),
                 + NCLASS,NCLS,MLINE,NCOL
 5                COMMON /CHAT/CONFUS(20,20)
                  INTEGER COL,SLINE,ELINE,RYLINE,SCOL,ECOL,RYCOL,TABLE
            C
            C
                  ILINE = 0
10                DO 100 I=1,SLINE
                  CALL GET LINE
                  ILINE = ILINE + 1
                100 CONTINUE
            C
15                DO 200 I=SLINE,ELINE,RYLINE
            C
                  DO 300 J=SCOL,ECOL,RYCOL
            C
                  DO 400 K=1,NCLASS
20                IF (COL(J).EQ.TABLE(K)) GO TO 500
                400 CONTINUE
                  GO TO 300
            C
                500 X = RANF(0)
25                DO 600 L=1,NCLASS
                  IF (X.LE.CONFUS(K,L)) GO TO 700
                600 CONTINUE
                  GO TO 300
            C
30              700 COL(J) = TABLE(L)
            C
                300 CONTINUE
            C
                  CALL PUT LINE
35          C
                  DO 800 M=1,RYLINE
                  ILINE = ILINE + 1
                  IF (ILINE.GT.ELINE) GO TO 800
                  CALL GET LINE
40              800 CONTINUE
            C
                200 CONTINUE
            C
                  RETURN
45                END
```

```
SUBROUTINE INFO       73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.33.23

1                     SUBROUTINE INFO (TYPE,S,E,BY,L,M,N,NU)
                      COMMON /INPUT/ ICARD(80)
                      COMMON /BUFR/ IA(1000),LENBUF
                      INTEGER S,E,BY,GETNUM
5               C
                C
                C
                100   CONTINUE
                      WRITE(6,600) TYPE
10              600   FORMAT(* PLEASE ENTER *,A4,* INFORMATION CONCERNING THE *
                     +*MAP-FILE  */)
                      READ(5,500) ICARD
                500   FORMAT(80A1)
                      IF (EOF(5).NE.0) GO TO 700
15                    I = 1
                      S = GET NUM (I)
                      E = GET NUM (I)
                      BY = GET NUM (I)
                      IF (S.EQ.0.OR.E.EQ.0.OR.BY.EQ.0) GO TO 200
20                    WRITE(6,605) S,E,BY
                605    FORMAT(1X,I7,2(3H , ,I7))
                         NU = FLOAT(E - S + 1)
                         NF = FLOAT(IA(L) - IA(M) + 1)/FLOAT(IA(N)) + .99
                           IF (NF.LT.NU) GO TO 250
25                           RETURN
                C
                C
                200   WRITE(6,610) TYPE
                610   FORMAT(* INCORRECT *,A4,* INPUT PARAMETER -RE ENTER-*/)
30                      GO TO 100
                C
                C
                250   WRITE(6,620) NF,NU
                620   FORMAT(* INPUT DIMENSIONS EXCEED MAP-FILE, MAP-FILE CONTAINS*/
35                   +1X,I6,* POINTS, YOUR ENTRY WAS *,I7,* POINTS*/)
                        GO TO 100
                C
                C
                700   WRITE(6,630)
40              630   FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT */)
                      RETURN
                      END
```

```
SUBROUTINE CLASIF     73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.33.23

 1                SUBROUTINE CLASIF
                  COMMON /INPUT/ ICARD(80)
                  COMMON /BUFR/ IA(1000),LENBUF
                  COMMON /DIMS/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,
 5                +TABLE(20),NCLASS,NCLS,NLINE,NCOL
                  COMMON /CMAT/ CONFUS(20,20)
                  INTEGER TABLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,GETNUM
            C
            C
10          C
                  NCLASS = NCLS/2
                  DO 11 I = 1,NCLASS
                  K = 2 * I + 8
                  TABLE(I) = IA(K)
15             11 CONTINUE
            C
                  DO 4 I2=1,NCLASS
                  DO 4 I1=1,NCLASS
                4 CONFUS(I1,I2) = 0.
20          C
                  WRITE(6,10)
                10 FORMAT(* DO YOU WANT TO USE AN EXISTING CONFUSION MATRIX (TAPE9)+*
                + )
                  READ(5,20)ANS
25              20 FORMAT(A1)
                  IF (ANS.NE.1HY) GO TO 100
                  READ (9) CONFUS
                  RETURN
            C
30           100 DO 200 I=1,NCLASS
                  K = 2*I + 1
                  WRITE(6,600) IA(K+8)
                 600 FORMAT(* ENTER CLASSIFICATION ACCURACY FOR CLASS *,A10/
                + * 10 ELEMENTS PER LINE, FORMAT(10F5.3)*)
35                WRITE(6,610)(IA(KK),KK=9,27,2)
                 610 FORMAT(1X,10(*/*,A3,*/*))
                 300 READ(5,500) (CONFUS(I,J),J=1,10)
                 500 FORMAT(10F5.3)
                  IF (EOF(5).NE.0) GO TO 700
40                IF (NCLASS.LE.10) GO TO 400
                  READ(5,500) (CONFUS(I,J),J=11,NCLASS)
                  IF (EOF(5).NE.0) GO TO 700
                 400 CONTINUE
                 200 CONTINUE
45          C
                  WRITE(6,60)
                  60 FORMAT(* DO YOU WANT TO STORE THE CONFUSION MATRIX (TAPE9)+*)
                  READ(5,20)ANS
                  IF (ANS.NE.1HY) RETURN
50                WRITE (9) CONFUS
                  RETURN
            C
                 700 WRITE(6,650)
                 650 FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT *)
55                RETURN
                  END
```

```
SUBROUTINE ACCUMC      73/73   OPT=1                    FTN 4.6+439        76/09/21. 14.33.23

 1                SUBROUTINE ACCUMC
                  COMMON /DIMS/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,TABLE(20),
                 + NCLASS,NCLS,NLINE,NCOL
                  COMMON /CMAT/ CONFUS(20,20)
 5          C
            C
                  DO 100 I=1,NCLASS
            C
                    DO 200 J=2,NCLASS
10                  JM1 = J - 1
                    CONFUS(I,J) = CONFUS(I,JM1) + CONFUS(I,J)
              200   CONTINUE
            C
              100 CONTINUE
15          C
                  RETURN
                  END
```

```
SUBROUTINE GETLINE    73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.33.23

 1                SUBROUTINE GET LINE
                  COMMON /BUFR/ A (1000),LENBUF
                  COMMON /DIMS/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,
                 +TABLE(20),NCLASS,NCLS,NLINE,NCOL
 5                EQUIVALENCE ( A(1),COL(1) )
                  INTEGER COL(1000),A
               C
               C
               C
10                BUFFER IN (7,1)  (A(1),A(1000))
                  IF (UNIT(7)) 111,770,770
                111  LENBUF = LENGTH(7)
                  N = LENBUF / 2
                  DECODE ( 10000,10,A(1) ) (COL(L),L=1,N)
15                10   FORMAT(500(A1,19X))
                  RETURN
               C
               C
               C
20                ENTRY HEADER
               C
               C
                  BUFFER IN (7,1) ( A(1),A(1000) )
                  IF (UNIT(7)) 110,770,770
25                110  N = LENGTH(7)
                  LENBUF = N
                  INCR = 15
                  IF (N.GE.100) INCR = 50
                  IK= 8 + INCR * 2
30                NCLS = 0
                  DO 200 J = 9,IK
                     IF (A(J).NE.0) NCLS = NCLS + 1
                200  CONTINUE
                  WRITE(6,600)(A(J),J=1,8),(A(K+8),K=1,NCLS)
35                600  FORMAT(*1   HEADER INFORMATION FROM MAP FILE  *,//,
                 +* NUMBER OF CHANNELS  *, I5,/,* NUMBER OF COLUMNS  *,I5,
                 +/,* LINE INFORMATION  *,I7,2(1H,,I7)/,
                 +* COLUMN INFORMATION  *,I7,2(1H,,I7),/,
                 +* CLASS NAMES AND SYMBOLS  *,/,(5(3X,A10,1X,A1)))
40                 N = A(1)
                  WRITE(6,610) (A(J+IK),J=1,N,1)
                610  FORMAT(/* CHANNELS REPRESENTED IN MAP-FILE */,(8(1X,I4))/)
                  BUFFER OUT (8,1) (A(1),A(LENBUF))
                  IF (UNIT(8)) 300,800,800
45                300  CONTINUE
                  RETURN
                770  WRITE(6,701)
                701  FORMAT(* END-OF-FILE ENCOUNTERED ON THE MAP-FILE *)
                  STOP 7000
50                800  WRITE(6,801)
                801  FORMAT(* ERROR IN BUFFER OUT *)
                  STOP 1111
                  END
```

```
FUNCTION GETNUM     73/73   OPT=1                    FTN 4.6+439        79/09/21. 14.33.23

 1                  INTEGER FUNCTION GET NUM (ICOL)
                    COMMON /INPUT/ ICARD(80)
                C
                C
 5                  GET NUM = 0
                C
                C
                    DO 100 I = ICOL , 80 , 1
                       IF (ICARD(I).EQ.55B) GO TO 100
10                     IF (ICARD(I).GE.33B.AND.ICARD(I).LE.44B) GO TO 200
                       IF (ICARD(I).GT.44B) GO TO 300
                 100 CONTINUE
                    ICOL = 80
                    RETURN
15              C
                C
                 200 GET NUM = GET NUM * 10 + ICARD(I) - 33B
                    GO TO 100
                 300 ICOL = I + 1
20                  RETURN
                       END
```

```
      SUBROUTINE PUTLINE    73/73   OPT=1                    FTN 4.6+439        79/09/21. 14.33.23

 1                SUBROUTINE PUT LINE
                  COMMON /BUFF/ A(1000),LENBUF
                  INTEGER BUF(1000),A
              C
 5            C
                  DO 100 I=1,LENBUF
              100 BUF(I) = 0
              C
                  ENCODE (10000,10,BUF(1)) (A(L),L=1,LENBUF)
10             10 FORMAT(500(A1,19X))
              C
                  BUFFER OUT (8,1) (BUF(1),BUF(LENBUF))
                  IF (UNIT(8)) 200,800,800
              200 RETURN
15            C
              800 WRITE(6,80)
               80 FORMAT(* TROUBLE ENCOUNTERED IN BUFFER OUT (TAPE8)*)
                  RETURN
              C
20                END
```

APPENDIX E

The Use of Variable Probability Sampling in Landsat Classification Accuracy Estimation

THE USE OF VARIABLE PROBABILITY SAMPLING IN LANDSAT CLASSIFICATION ACCURACY ESTIMATION

An experiment proposed in the original study plan concerned the use of variable probability sampling in Landsat classification accuracy estimation. Although a method of using variable probability sampling for accuracy estimation was devised and a computer program to be used in this method written, a detailed analysis of the procedure was not performed. The two major reasons for this are: precise registration between the Landsat classification and the topographic base map was not obtained; and, a complete set of ground truth data with which to test the effectiveness of the procedure was not obtained. Although these two factors did not hinder the acreage estimation experiments, the proposed classification accuracy estimation procedure requires precise registration to be effective.

The method involves a random selection of sample units of a specified number of pixels on the Landsat classification. The selected sample units are then compared, pixel by pixel, with a ground truth map to determine the classification accuracy of the sample unit. These figures for sample unit classification accuracy are then plugged into the sampling formula to obtain an accuracy estimate for the entire classification.

As with the acreage estimation samples, the selection probabilities of sample units are weighted. In this case, the weight consists of an initial estimate of classification accuracy; training and/or test field classification accuracy. It gives those sample units estimated to be the least accuracy the greatest selection probability.

In the acreage estimation samples, the Landsat classification was gridded off into sample units, the number of pixels in a sample unit of the class of interest were counted, and this count was accumulated over all possible sample units in the population. In effect, each pixel is given the value one if it is of the class of interest and zero otherwise. These values are summed for each sample unit and for the entire population. In the accuracy estimation sample, each pixel is given a value between zero and one depending on the training or test field classification accuracy of the class of that pixel. Pixels of those classes of less training or test set classification accuracy will be assigned a higher value. The value assigned each pixel is equal to one minus the training or test set classification accuracy expressed as a decimal fraction.

As an example, consider the hypothetical classification of Figure 21, consisting of three land cover classes. The training set classification accuracy of each class is given in Table 40. Then the value assigned to each forest pixel = 1 - .85 = .15. The weight assigned to each grass pixel = 1 - .70 = .30. The weight assigned to each water pixel = 1 - .99 = .01.

The procedures to determine sample unit size, sample size, and to perform the list sample are the same as for the acreage estimation samples. Assume that a sample unit size of two pixels by two pixels has been selected. If the classification is gridded off into sample units, the appropriate weight applied to each pixel, and this value accumulated for each sample unit and for the population, the sample list of Table 41 is obtained.

Assume that two sample units are to be included in the sample. Two random numbers between 0.00 and 2.59 must then be drawn. If the two random numbers are 1.42 and 2.03 then sample units 2 and 3 are included in the sample. Ground truth information is then used to determine the number of misclassified pixels in each sample unit, y_i. Then an estimate of the number of misclassified pixels in the entire population is given by:

$$\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{p_i} \tag{3.6}$$

Suppose that in sample unit 2, two pixels are misclassified. Then $y_i = 2$. If sample unit 3 has one pixel misclassified, then $y_i = 1$. Then an estimate of misclassified pixels in the entire classification is:

$$\hat{A} = \frac{1}{2} \sum_{i=1}^{2} \frac{y_i}{p_i} = \frac{1}{2} \left(\frac{2}{.405} + \frac{1}{.178}\right) = 5.3$$

The sample standard deviation is:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{y_i}{p_i} - \hat{A}\right)^2 \tag{3.7}$$

$$= \frac{1}{n-1} \left[\sum_{i=1}^{n} \left(\frac{y_i}{p_i}\right)^2 - \frac{1}{n} \left(\sum_{i=1}^{n} \frac{y_i}{p_i}\right)^2\right] \tag{3.8}$$

$$= \frac{1}{1} \left[55.9 - \frac{1}{2} (111.4)\right]$$

$$= .23$$

$$s = \sqrt{s^2} \tag{3.9}$$

$$= \sqrt{.23} = .48$$

Thus, an estimate of the number of misclassified pixels in the entire population is 5.3. This gives an estimate of classification

accuracy of (16 - 5.3)/16 = 66.9%. A 90% confidence interval for the estimate of 5.3 misclassified pixels is given by:

$$\hat{A} \pm t\, s / \sqrt{n}$$

$$= 5.3 \pm (6.314)\ (.48) / \sqrt{2}$$

$$= 5.3 \pm 2.1 = (3.2,\ 7.4)$$

A 90% confidence interval for classification accuracy is then given by:

$$[(16 - 7.4)/16,\quad (16 - 3.2)/16]$$

$$= (53.8\%,\ 80.0\%)$$

A computer program, PPSACC, was written to apply the correct weight to each pixel of a Landsat classification, grid off the classification into sample units of a specified size and shape, and accumulate the weighted values for each sample unit and for the entire population. In addition to the Landsat classification input to the program, a list of weights, corresponding to the training or test field classification accuracy (or any other rough estimate of accuracy) for each class is input. The output is a sample list similar to Table 7 which can then be used in the list sampling procedure. This program is a modification of the program PSU, and produces a sample list in the same format as PSU.

PPSACC was run for the Middle Park study area using the ORIGINAL classification. Sample units were 25 pixels by 25 pixels. The weights for each class were the diagonal elements of the training field accuracy confusion matrix of Table 20. These weights are equal to one minus the percentage of correctly classified pixels (expressed as a decimal fraction). The weights used, including an arbitrary value assigned to the non-vegetated/disturbed class, are given in Table 42.

The sample list output by PPSACC is shown in Table 43. To select the specific sample units to include in the sample, list sampling is used. Random numbers between 0 and 4626.370 are drawn, and the appropriate sample units selected. Those sample units are then compared to ground truth data to determine the number of misclassified pixels. Formulas 3.6 through 3.9 are used to obtain estimates of number of misclassified pixels in the entire classification and sample standard deviation. In this case, the value y_i is equal to the number of misclassified pixels in the sample unit i; x_i is obtained from the sample list; p_i is equal to $x_i \; / \; \sum_{i=1}^{N} x_i = x_i \; / \; 4626.370$. An estimate of total classification accuracy is given by:

$$(m - \hat{A})/m$$

where:

m = total number of pixels in the classification

$\hat{A}$ = estimate of number of misclassified pixels

As previously mentioned, an actual experiment was not conducted. It is expected that the advantage of using variable probability sampling rather than simple random sampling would be the same as that seen for acreage estimation, namely a decrease in sample size needed to achieve a specified degree of confidence.

Two differences should be noted, however, between the use of this sampling method for acreage estimation and classification accuracy estimation. First, the values of x_i, as seen in the sample list, show much less variability for accuracy estimation than for acreage estimation. This could tend to decrease the advantages of using variable probability sampling versus simple random sampling. Second, the training

field accuracy estimates used to determine selection probabilities are very rough estimates, which may not be very well correlated with actual classification accuracy. The only way to test the effectiveness of this procedure would be conduct an actual experiment and compare the results to other methods of accuracy estimation. It is believed, however, that the most important factors in any method used are the need for high quality gound truth data and accurate registration between the Landsat image and the ground truth maps.

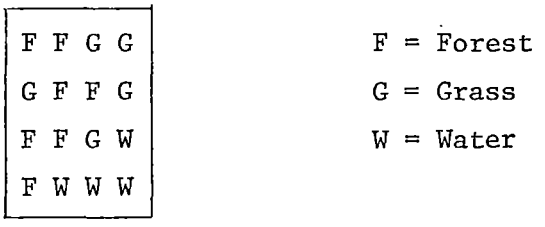


Figure 21. Hypothetical Landsat classification.

Table 40. Training set classification accuracy.

Forest	85%
Grass	70%
Water	99%

Table 41. Sample list.

Sample Unit	Lines	Cols	x_i	$\sum_{j=1}^{i} x_j$	$p_i = x_i / \sum_{i=1}^{n} x_i$
1	1-2	1-2	.15+.15+.30+.15=0.75	0.75	0.75/2.59=0.290
2	1-2	3-4	.30+.30+.15+.30=1.05	1.80	1.05/2.59=0.405
3	3-4	1-2	.15+.15+.15+.01=0.46	2.26	0.46/2.59=0.178
4	3-4	3-4	.30+.01+.01+.01=0.33	2.59	0.33/2.59=0.127

Table 42. Class accuracy weights input to PPSACC.

Class	Weight
Deciduous	.075
Coniferous	.069
Non-vegetated/disturbed	.100
Mountain grass	.056
Meadow	.273
Mountain shrub	.167
Water	.008
Sage	.091

Table 43. Sample list.

PRIMARY SAMPLE UNITS FOR			ASPEN PASTR		ASPN SHRUB	CONIF WATER		DISTR YSAGE	GRASS YPERL
PSU #	LINES			COLS.			Y(I)	SUM Y(I)	
1	1	25	1	1	25	1	74.9	74.9	
2	1	25	1	26	50	1	68.0	142.9	
3	1	25	1	51	75	1	64.2	207.1	
4	1	25	1	76	100	1	109.0	316.1	
5	1	25	1	101	125	1	67.4	383.4	
6	1	25	1	126	150	1	81.2	464.6	
7	1	25	1	151	175	1	70.7	535.3	
8	1	25	1	176	200	1	68.9	604.2	
9	26	50	1	1	25	1	75.9	680.1	
10	26	50	1	26	50	1	74.4	754.5	
11	26	50	1	51	75	1	57.6	812.1	
12	26	50	1	76	100	1	86.2	898.2	
13	26	50	1	101	125	1	71.7	969.9	
14	26	50	1	126	150	1	68.0	1037.9	
15	26	50	1	151	175	1	71.2	1109.1	
16	26	50	1	176	200	1	77.5	1186.6	
17	51	75	1	1	25	1	88.3	1274.9	
18	51	75	1	26	50	1	89.0	1363.9	
19	51	75	1	51	75	1	70.3	1434.2	
20	51	75	1	76	100	1	76.3	1510.5	
21	51	75	1	101	125	1	87.1	1597.6	
22	51	75	1	126	150	1	55.5	1653.1	
23	51	75	1	151	175	1	66.5	1719.6	
24	51	75	1	176	200	1	69.2	1788.9	
25	76	100	1	1	25	1	91.8	1880.7	
26	76	100	1	26	50	1	93.2	1973.9	
27	76	100	1	51	75	1	66.4	2040.3	
28	76	100	1	76	100	1	66.6	2106.9	
29	76	100	1	101	125	1	75.1	2182.0	
30	76	100	1	126	150	1	77.9	2259.9	
31	76	100	1	151	175	1	83.8	2343.7	
32	76	100	1	176	200	1	81.2	2424.9	
33	101	125	1	1	25	1	61.6	2486.5	
34	101	125	1	26	50	1	93.0	2579.5	
35	101	125	1	51	75	1	64.4	2643.9	
36	101	125	1	76	100	1	62.1	2705.9	
37	101	125	1	101	125	1	53.0	2758.9	
38	101	125	1	126	150	1	83.7	2842.7	
39	101	125	1	151	175	1	85.7	2928.3	
40	101	125	1	176	200	1	62.7	2991.0	
41	126	150	1	1	25	1	63.0	3054.0	
42	126	150	1	26	50	1	85.0	3139.0	
43	126	150	1	51	75	1	77.2	3216.2	
44	126	150	1	76	100	1	59.0	3275.2	
45	126	150	1	101	125	1	68.1	3343.3	
46	126	150	1	126	150	1	50.0	3393.3	
47	126	150	1	151	175	1	71.4	3464.7	
48	126	150	1	176	200	1	56.3	3521.0	
49	151	175	1	1	25	1	64.9	3586.0	
50	151	175	1	26	50	1	62.3	3648.3	
51	151	175	1	51	75	1	73.1	3721.4	
52	151	175	1	76	100	1	84.1	3805.5	
53	151	175	1	101	125	1	75.5	3880.9	
54	151	175	1	126	150	1	64.6	3945.5	
55	151	175	1	151	175	1	50.9	3996.4	
56	151	175	1	176	200	1	57.2	4053.6	
57	176	200	1	1	25	1	82.8	4136.3	
58	176	200	1	26	50	1	86.7	4223.0	
59	176	200	1	51	75	1	56.0	4279.1	
60	176	200	1	76	100	1	60.0	4339.1	
61	176	200	1	101	125	1	61.1	4400.1	
62	176	200	1	126	150	1	87.4	4487.5	
63	176	200	1	151	175	1	75.2	4562.7	
64	176	200	1	176	200	1	63.6	4626.3	

Table 43 (continued).

MEAN = 72.28703
STANDARD DEVIATION = 12.07815
COEFFICIENT OF VARIATION = .16709

```
      SUBROUTINE GETLINE    73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                SUBROUTINE GET LINE
                  COMMON /BUF/ A (500)
                  COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
                 +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
 5                EQUIVALENCE ( A(1),COL(1) )
                  INTEGER COL(250),A
               C
               C
               C
10                BUFFER IN (7,1)  (A(1),A(500))
                  IF (UNIT(7)) 111,770,770
                111  N = LENGTH(7) / 2
                  DECODE ( 5000,10,A(1) ) (COL(L),L=1,N)
                 10   FORMAT(250(A1,19X))
15                RETURN
               C
               C
               C
                  ENTRY HEADER
20             C
               C
                  BUFFER IN (7,1) ( A(1),A(500) )
                  IF (UNIT(7)) 110,770,770
                110  N = LENGTH(7)
25                INCR = 15
                  IF (N.GE.100) INCR = 50
                  IK= 8 + INCR * 2
                  NCLS = 0
                  DO 200 J = 9,IK
30                   IF (A(J).NE.0) NCLS = NCLS + 1
                 200  CONTINUE
                  WRITE(8,600)(A(J),J=1,8),(A(K+8),K=1,NCLS)
                  WRITE(6,600)(A(J),J=1,8),(A(K+8),K=1,NCLS)
                 600  FORMAT(*1   HEADER INFORMATION FROM MAP FILE  *,//,
35                +* NUMBER OF CHANNELS  *, I5,/,* NUMBER OF COLUMNS  *,I5,
                  +/,* LINE INFORMATION  *,I7,2(1H,,I7)/,
                  +* COLUMN INFORMATION  *,I7,2(1H,,I7),/,
                  +* CLASS NAMES AND SYMBOLS  *,/,(5(3X,A10,1X,A1)))
                    N = A(1)
40                WRITE(8,610) (A(J+IK),J=1,N,1)
                  WRITE(6,610) (A(J+IK),J=1,N,1)
                 610  FORMAT(/* CHANNELS REPRESENTED IN MAP-FILE */,(8(1X,I4))/)
                  RETURN
                 770  WRITE(6,701)
45               701  FORMAT(* END-OF-FILE ENCOUNTERED ON THE MAP-FILE *)
                  STOP 7000
                  END
```

```
      SUBROUTINE STATS      73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                SUBROUTINE STATS
                  COMMON /STATS/ SUM,SQM,IPSU
               C
               C
 5                XMEAN = SUM / FLOAT(IPSU)
                  IF (IPSU.LE.1) STOP 7777
                  STD = SQRT (( SQM-(SUM**2/FLOAT(IPSU))) / FLOAT (IPSU-1) )
                  CV = STD / XMEAN
                  WRITE(6,600) XMEAN,STD,CV
10                WRITE(8,600) XMEAN,STD,CV
                 600  FORMAT(*1*,5X,*MEAN = *,F15.5,/5X,*STANDARD DEVIATION = *,
                +F15.5,/5X,*COEFFICIENT OF VARIATION = *,F15.5)
                  RETURN
                  END

10.32.34.UCLP, 36,      1.728KLNS.      FAIJFIP
```

```
                                                                    FTN 4.6+439          78/09/21. 14.36.11

 1                SUBROUTINE QUIZ
                  COMMON /INPUT/ CARD(80),SCREEN
                  LOGICAL SCREEN
                C
 5              C
                  SCREEN = .F.
                  WRITE(6,600)
                 600  FORMAT(1X,*WOULD YOU LIKE A DETAILED LISTING OF EACH PSU **)
                  READ(5,500) CARD
10                 IF (EOF(5).NE.0) GO TO 770
                 500  FORMAT(80R1)
                  I = 1
                   IF ( NAME(I).EQ.3HYES) SCREEN = .TRUE.
                 770  RETURN
15                END
```

```
 SUBROUTINE OUTPUT      73/73   OPT=1                    FTN 4.6+439          78/09/21. 14.36.11

 1                SUBROUTINE OUTPUT ( ICOUNT,LINE )
                  COMMON /INPUT/ CARD(80),SCREEN
                  COMMON /STATS/ SUM,SQM,IPSU
                  COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
 5               +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                  INTEGER BYCOL,BYLINE,SCOL,SLINE
                  LOGICAL SCREEN
                C
                C
10              C
                  LINE = LINE + NBINS
                  IF ( LINE.LE.60 ) GO TO 200
                     IF ( SCREEN ) WRITE(6,599) (NAMES(J),J=1,NCLASS)
                     IF ( SCREEN ) WRITE(6,600)
15              C
                C
                  WRITE(8,599)  (NAMES(J),J=1,NCLASS)
                 599  FORMAT(*1*,5X,*PRIMARY SAMPLE UNITS FOR   *,5A10,/,33X,
                +5A10,/)
                  WRITE(8,600)
20               600  FORMAT (5X,*PSU #*,11X,*LINES*,17X,*COLS.*,11X,*X(I) *,
                +5X,* SUM  X(I)*,/)
                  LINE = 0
                C
25              C
                C
                 200  IM =  ILENGHT
                  IN =  IWIDTH
                  IFL = SLINE + ( ICOUNT * IM ) - IM
30                ILL = IM - BYLINE  + IFL
                  IFC = SCOL - IN
                C
                C
                C
35                DO 300 I = 1 , NBINS
                     SUM = SUM + BIN(I)
                     SQM = SQM + BIN(I)**2
                     IPSU = IPSU + 1
                     IFC = IFC + IN
40                   ILC = IFC + IN - BYCOL
                        IF ( SCREEN ) WRITE(6,660)IPSU,IFL,ILL,BYLINE,IFC,ILC
                +,BYCOL,BIN(I),SUM
                        WRITE(8,660) IPSU,IFL,ILL,BYLINE,IFC,ILC,BYCOL,BIN(I),SUM
                 660  FORMAT(5X,I5,5X,I5,1X,I5,1X,I5,5X,I5,1X,I5,1X,I5,5X,F5.1,5X,F10.3)
45               300  CONTINUE
                  RETURN
                  END
```

```
SUBROUTINE CLASIF      73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                SUBROUTINE CLASIF
                  COMMON /INPUT/ ICARD(80),SCREEN
                  COMMON /BUF/ IA(500)
                  COMMON /DIM/ SLINE,FLINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
 5               +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                  COMMON /WTS/ WEIGHT(20)
                  INTEGER TABLE,SLINE,FLINE,BYLINE,SCOL,ECOL,BYCOL,BIN,GETNUM
               C
               C
10             C  ***************************************************************************
               C  FOLLOWING ROUTINE CHANGED FOR PPSACC MODIFICATIONS
               C  ***************************************************************************
                     NCLASS = NCLS/2
                  100  DO 200 I=1,NCLASS
15                    K = 2*I - 1
                      WRITE(6,600) IA(K+8)
                  600  FORMAT(* PLEASE ENTER WEIGHT FOR CLASS *,A10/)
                      READ(5,500) WEIGHT(I)
                  500  FORMAT(F7.5)
20                    IF (EOF(5).NE.0) GO TO 700
                      NAMES(I) = IA(K+8)
                      TABLE(I) = IA(K+9)
                  200  CONTINUE
                      RETURN
25             C  ***************************************************************************
               C  ***************************************************************************
                  700  WRITE(6,650)
                  650  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT *)
                      RETURN
30                    END
```

```
SUBROUTINE PSUSIZE    73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

      1               SUBROUTINE PSU SIZE
                      COMMON /INPUT/ ICARD(80),SCREEN
                      COMMON /BUF/ IA(500)
                      COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,IPSUL,IPSUC,
      5              +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                      INTEGER TAPLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,BIN,GETNUM
                C
                C
                C
     10               100  WRITE(6,600)
                      600  FORMAT(* PLEASE ENTER THE PSU DIMENSIONS */)
                           READ(5,500) ICARD
                      500  FORMAT(80R1)
                           IF (EOF(5).NE.0) GO TO 700
     15                        I = 1
                               IPSUL = GET NUM (I)
                               IPSUC = GET NUM (I)
                               IF(IPSUL.EQ.0.OR.IPSUC.EQ.0) GO TO 200
                               WRITE(6,605) IPSUL,IPSUC
     20                605          FORMAT(1X,I7,3H , I7)
                                    IF(MOD(NLINE,IPSUL).NE.0) GO TO 250
                                    IF(MOD(NCOL,IPSUC).NE.0) GO TO 250
                                    NBINS = NCOL/IPSUC
                                         RETURN
     25                C
                       C
                        200  WRITE(6,610)
                        610  FORMAT(* INCORRECT DIMENSIONS FOR PSU -RE ENTER-*/)
                                    GO TO 100
     30                C
                       C
                        250  WRITE(6,620)
                       620   FORMAT(* PSU DIMENSIONS AS ENTERED ARE NOT AN INTEGRAL *,/,1X,
                            +*NUMBER OF THE FILE DIMENSIONS PREVIOUSLY INPUT*/)
     35                          GO TO 100
                        700  WRITE(6,607)
                        607  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT *)
                              STOP 6000
                           END
```

```
SUBROUTINE INFO       73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                SUBROUTINE INFO (TYPE,S,E,BY,L,M,N,NU)
                  COMMON /INPUT/ ICARD(80),SCREEN
                  COMMON /BUF/ IA(500)
                  INTEGER S,E,BY,GETNUM
 5          C
            C
            C
             100  CONTINUE
                  WRITE(6,600) TYPE
10           600  FORMAT(* PLEASE ENTER *,A4,* INFORMATION CONCERNING THE *
                 +*MAP-FILE  */)
                  READ(5,500) ICARD
             500  FORMAT(80R1)
                  IF (EOF(5).NE.0) GO TO 700
15                I = 1
                  S = GET NUM (I)
                  E = GET NUM (I)
                  BY = GET NUM (I)
                  IF (S.EQ.0.OR.E.EQ.0.OR.BY.EQ.0) GO TO 200
20                WRITE(6,605) S,E,BY
             605   FORMAT(1X,I7,2(3H , ,I7))
                     NU = FLOAT(E - S + 1)
                     NF = FLOAT(IA(L) - IA(M) + 1)/FLOAT(IA(N)) + .99
                       IF (NF.LT.NU) GO TO 250
25                       RETURN
            C
            C
             200  WRITE(6,610) TYPE
             610  FORMAT(* INCORRECT *,A4,* INPUT PARAMETER -RE ENTER-*/)
30                 GO TO 100
            C
            C
             250  WRITE(6,620) NF,NU
             620  FORMAT(* INPUT DIMENSIONS EXCEED MAP-FILE, MAP-FILE CONTAINS*/
35                +1X,I6,* POINTS, YOUR ENTRY WAS *,I7,* POINTS*/)
                    GO TO 100
            C
            C
             700  WRITE(6,630)
40           630  FORMAT(* END-OF-FILE ENCOUNTERED ON INPUT */)
                  RETURN
                  END
```

SUBROUTINE CORE

```
      SUBROUTINE CORE
      COMMON /STATS/ SUM,SQM,IPSU
      COMMON /BUF/ A (500)
      COMMON /DIM/ SLINE,FLINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
     +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
      COMMON /WTS/ WEIGHT(20)
      EQUIVALENCE (A(1),COL(1))
      INTEGER COL(500),SLINE,FLINE,BYLINE,SCOL,ECOL,BYCOL,TABLE
C
C
      ICOUNT = 0
      SUM = SQM = 0.0
      IPSU = 0
      K = SLINE
      ILINE = 0
      LINE = 100
C
C
      DO 200 I = 1,NBINS
          BIN(I) = 0.0
  200 CONTINUE
C
C
      DO 300 I = SLINE , FLINE , BYLINE
C
C
        DO 301 J = 1 , K
           CALL GET LINE
  301    CONTINUE
         K = BYLINE
           DO 302 L = SCOL , ECOL , BYCOL
C
C
              DO 303 M = 1 , NCLASS
                 IF ( COL(L).EQ.TABLE(M) ) GO TO 304
  303         CONTINUE
                    GO TO 302
  304          IBIN = NBINS - ( ECOL - L ) / (  IWIDTH )
C  ***************************************************************************
C  FOLLOWING ROUTINE CHANGED FOR PPSACC MODIFICATIONS
C  ***************************************************************************
      BIN(IBIN) = BIN(IBIN) + WEIGHT(M)
C  ***************************************************************************
C  ***************************************************************************
  302    CONTINUE
C
C
         ILINE = ILINE + BYLINE
         IF ( ILINE.NE.ILENGHT ) GO TO 300
            ILINE = 0
            ICOUNT = ICOUNT + 1
            CALL OUTPUT ( ICOUNT,LINE )
C
C
              DO 310 II = 1,NBINS
                 BIN(II) = 0.0
  310         CONTINUE
C
  300 CONTINUE
C
C
      CALL STATS
      RETURN
      END
```

```
    FUNCTION GETNUM      73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                INTEGER FUNCTION GET NUM (ICOL)
                  COMMON /INPUT/ ICARD(80),SCREEN
            C
            C
 5                GET NUM = 0
            C
            C
                  DO 100 I = ICOL , 80 , 1
                      IF (ICARD(I).EQ.55B) GO TO 100
10                    IF (ICARD(I).GE.33B.AND.ICARD(I).LE.44B) GO TO 200
                      IF (ICARD(I).GT.44B) GO TO 300
             100  CONTINUE
                  ICOL = 80
                  RETURN
15          C
            C
             200  GET NUM = GET NUM * 10 + ICARD(I) - 33B
                  GO TO 100
             300  ICOL = I + 1
20                RETURN
                      END
```

```
    FUNCTION NAME        73/73   OPT=1                    FTN 4.6+439        78/09/21. 14.36.11

 1                INTEGER FUNCTION NAME (ICOL)
                  COMMON /INPUT/ ICARD(80),SCREEN
            C
            C
 5          C
                  NAME = 10H
                  ICHAR = 0
                  DO 100 I = ICOL , 80 , 1
            C
10                    IF(ICARD(I).EQ.55B) GO TO 100
                      IF(ICARD(I).GT.0.AND.ICARD(I).LE.44B) GO TO 200
                      IF(ICARD(I).GT.44B) GO TO 300
             100   CONTINUE
            C
15                I = 79
             300  ICOL = I + 1
                  J = (10 - ICHAR) * 6
                  NAME = SHIFT(NAME,J)
                  RETURN
20          C
            C
            C
             200  ICHAR = ICHAR + 1
                  NAME = SHIFT (NAME,6).AND.77777777777777777008
25                NAME = NAME .OR. ICARD(I)
                  GO TO 100
                  END
```

```
      PROGRAM PPSACC      73/73   OPT=1                FTN 4.6+439        78/09/21. 14.36.11

  1                 PROGRAM PPSACC(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT,TAPE7,TAPE8)
                    COMMON /INPUT/ ICARD(80),SCREEN
                    COMMON /BUF/ IA(500)
                    COMMON /DIM/ SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,ILENGHT,IWIDTH,
  5                +TABLE(20),NAMES(20),NCLASS,NCLS,NBINS,NLINE,NCOL,BIN(250),INCR
                    COMMON /WTS/ WEIGHT(20)
                C
                C
                C
 10                 INTEGER TABLE,SLINE,ELINE,BYLINE,SCOL,ECOL,BYCOL,BIN,GETNUM
                C
                C
                C
                    CALL HEADER
 15                 CALL INFO (4HLINE,SLINE,ELINE,BYLINE,4,3,5,NLINE)
                    CALL INFO (4HCOL ,SCOL,ECOL,BYCOL,7,6,8,NCOL)
                    CALL PSU SIZE
                    CALL CLASIF
                    CALL QUIZ
 20                 CALL CORE
                    END
```